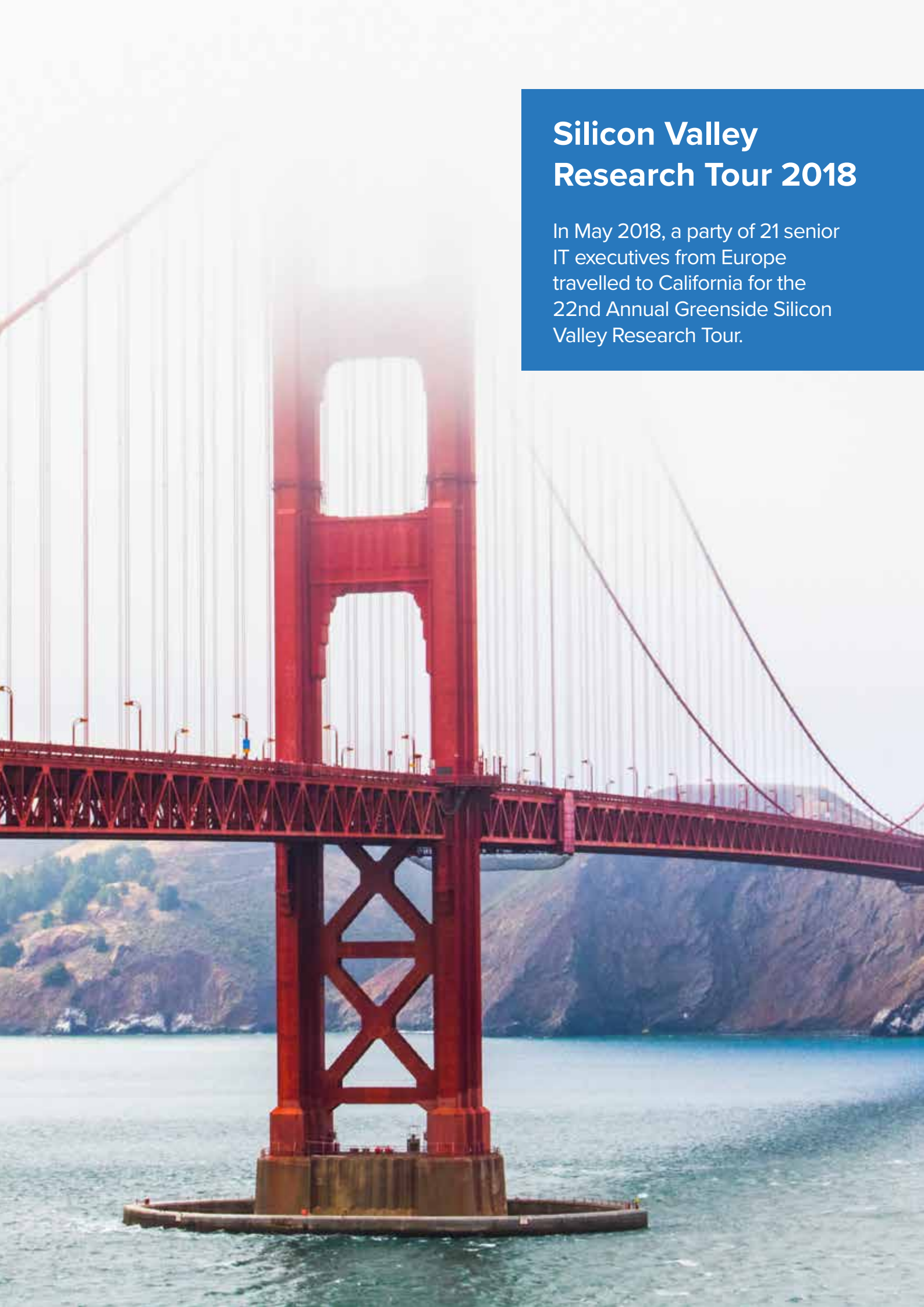
Silicon Valley Research Tour 2018

Ntegra Greenside

Knowledge for IT Leaders



Silicon Valley Research Tour 2018

In May 2018, a party of 21 senior IT executives from Europe travelled to California for the 22nd Annual Greenside Silicon Valley Research Tour.

Company Index

Executive Introduction

In May 2018, a party of 21 senior IT executives from Europe travelled to California for the 22nd Annual Greenside Silicon Valley Research Tour. This was the first US Tour to be run exclusively by Ntegra and we were keen make some changes as well as continue to build on the foundations established by Mike and Linda Portlock. Our aim was to increase the value the Tour provides to both Greenside alumni and new delegates. The changes included moving our base from The Fairmont Hotel, San Jose to the Stanford Park Hotel in Menlo Park which significantly reduced travel time and provided better access to facilities in Palo Alto and the Stanford Shopping Mall, both within easy walking distance from the hotel.

During the weeklong tour, we received presentations from Venture Capitalists and a carefully curated group of innovative start-up and scale-up technology companies. On Wednesday we stayed a full day at the hotel where we held our first panel discussion with executives from VCs, start-up founders and entrepreneurs.

The purpose of the tour was to investigate digital enablement and learn as much as possible about emerging developments in the deployment and application of exciting new information technology innovations. We received presentations from 35 companies of varying sizes and maturity which were actively involved in providing products and services that help organisations to become more effective in their use and exploitation of technology. In addition, we received presentations from host Venture Capital firms, Andreessen Horowitz, Sequoia Capital, Next World Capital and Greylock Partners.

Silicon Valley remains a unique and fascinating place to visit. With its own culture and methods, there's nowhere else quite like it. The blend of leadership, talent and access to funding is second to none. The scale of the Valley outstrips anywhere else in the world by an order of magnitude; Silicon Valley indexes higher than any other ecosystem in all measures (global research, access to funding, performance and talent). The Ntegra Greenside programme provides a unique conduit between this ecosystem and enterprise technology leaders from the UK and Europe.

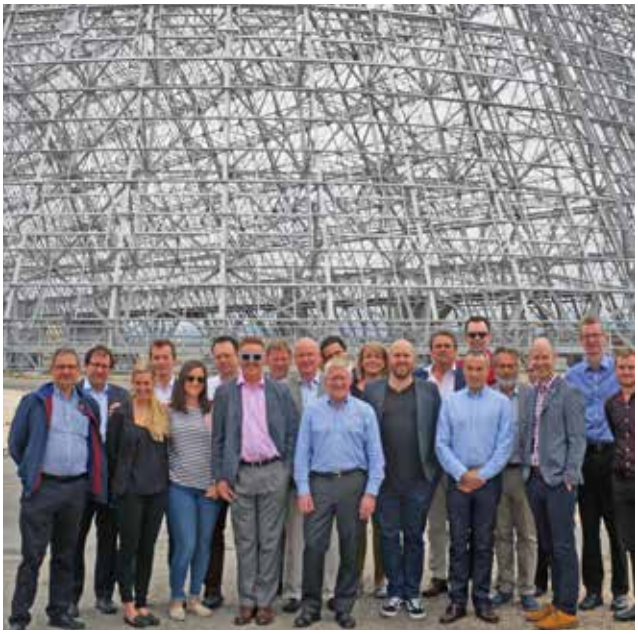
As in previous years, there were some strong recurrent themes, with domains reaching new levels of maturity, such as: Data and Analytics; the Cyber arms-race; and Artificial Intelligence and Machine Learning. Other, previously emerging themes have now just become accepted as the norm: velocity of change; cloud first; open source; everything (as a service); DevOps and software defined infrastructure (including VM, containers, server-less, etc.). New themes this year included: the hunt for the next big wave in core platforms; the challenge of simply keeping up; cloud/Internet scale at low cost; and the cognitive enterprise.

While visiting Andreessen Horowitz at the beginning of the week, the tour received an update from Benedict Evan's on his "10 Year Futures" view. His insights were closely echoed by the panel members during our midweek session and by Jeff Pashalides in his closing presentation at Sequoia. Silicon Valley is fuelled by big platform transitions, from silicon chips to PCs, from to LANs to WANs, to the Internet fuelling telecommunications and now mobile (which is still "Eating the World" but, perhaps, with improved table manners). What we are talking about in innovation conversations is what's coming next? Will it be new capabilities supported by the billion+ unit platforms, or will it be a new a new platform adoption s-curve?

Fast forward 2019 and what big platform movement might we see? Will it be in Autonomy, which is still finding its feet and not quite working yet? Mixed reality, which is starting to work well. Crypto currency, underpinned by the blockchain, which is proven and works very well but no one can figure out what the market really is yet. What seems most likely is AI, which everyone in the Valley and further afield are pouring rocket-fuel on. Our Wednesday keynote speaker this year, Jim Spohrer (Director of Cognitive OpenTech at IBM), provided an excellent and thought-provoking presentation that supports our view that AI and IA development and continuing maturity will dominate the next year.

Planning is already underway for our 23rd US tour in 2019, with many of these themes and topics expected to play a major role. I hope to see you there in 2019!

Andy Jefferies

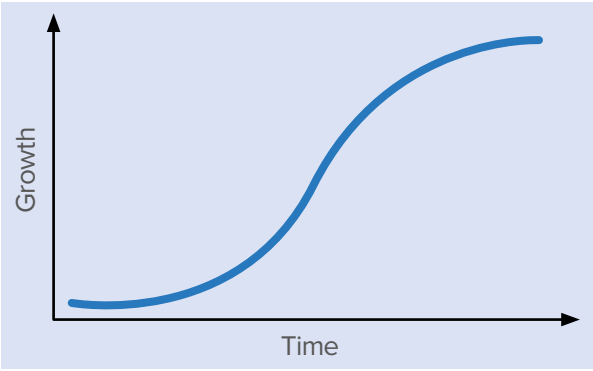


P.S. This year we are delighted to announce our first Israel Research Tour which will deliver a different and fresh perspective to our clients and fits nicely into the annual research calendar. With the US Research Tour completing its 22nd successful year, and our European Research Tours running for 5 years, we've introduced tour delegates to hundreds of technology companies, many of whom have gone on to become major disruptors, transforming businesses around the world. Israel ranks 4th globally in scientific research (measured via number of publications) with expenditure on technology and scientific R&D being amongst the highest in the world. It has the highest number of scientists and technologists per capita of any country in the world (140 per 10K employees, compared to 88 per 10K employee in the US). Israel has the second largest start-up scenes and supporting ecosystem in the world, again second only to Silicon Valley.

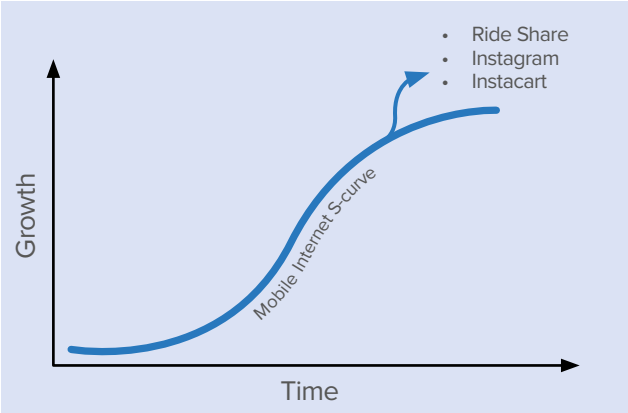
Benedict Evans

Benedict Evans provided an update on his "10 Year Futures" research with his views on "what's happening now". He talked about the state of technology today and what's likely to happen in the next decade: mobile, Google / Apple / Facebook / Amazon, innovation, machine learning, autonomous cars, mixed reality and crypto-currencies. His thinking is a good example of both 'S-curves' and 'Three Horizons' thinking and how the S-curve life-cycle can be summed-up by saying "The winners always look invulnerable, until they don't!"

Benedict started by introducing the S-Curve that's dominated and created the technology industry in the last few decades, the growth of the PC, which went from zero to around 1.5 billion units installed around the world in 20 or so years. This has been replaced by another S-curve, smartphones, which in turn has led to the growth in mobile-phones, which connect just about everyone on the planet. He says, "There are maybe ~3 billion smartphones and 5 billion mobile-phones connecting about 5.5 billion adults around the world." What we see with all these S-curves is that growth starts slowly and then "explodes" before calming down again and becoming "a little bit boring."



When we start to systematise this we see S-curves following S-curves. Most recently we had the PC Internet which was followed by the Mobile Internet S-curve. When we get to the stable and more mature part at the end of each S-curve, we start thinking about what we can build on top of it. For example, on the PC internet S-curve we talked about building Search and Social. Now on top of the MobileInternet S-curve we talk about ride-sharing, Instagram and Instacart.



In talking about what can be built on top of the platform rather than the platform itself, we start to induce a feeling that things are slowing down, becoming boring and that "innovation is dead," "the new iPhone looks just like the old iPhone," and "what have you done for me lately?" Benedict says "of course, it's always like this." He then referred to a young graduate named Marc Andreessen who said, "I arrived in Silicon Valley in 1994 and I thought I was too late and missed the whole thing." That wasn't quite right because there's always something new coming.

Today, we're having 2 sets of innovation conversations:

- What can we build on the billion scale platforms we have today?
- What are the next S-curves; what will the next billion scale platforms be?

All this conjecture is against a background where there are massive amounts of experimentation around the current platforms. New models and endpoints such as:

- Bots, Voice, Smart Speakers, Wearables, Watches and Smart Homes
- Where do TV shows fit and how do they align with marketing and lock-in?
- Social as pop culture, cameras as platforms
- VR? AR? Accessory or next platform?

There is also a 'meta-conversation' about "Is there any white-space left?" and is it possible to "compete with Google/Apple/Facebook/Amazon?"

There is a thesis emerging in Silicon Valley that the new super-evolved GAFA organisms seem to have a different character than the huge technology companies that went before them. The four key factors:

- Massively greater scale
- Everyone "read the disruption book" and saw Yahoo and MySpace fail
- Founder controlled (except Apple)
- There are 4, not 1

On the other hand, there have always been companies that look invulnerable, "until they don't." Think IBM then Microsoft, Intel and Nokia, then AOL, MySpace and Yahoo. No one could predict how and why they would fall and fail, but they all did, so the question remains what will happen to the GAFA companies and who or what will come thereafter?

Although, "Big scary tech giants" such as IBM have been around for a long time and have been proven to be non-invincible, Benedict postulates there may be something different about the GAFA companies. He presented the annual revenue of the GAFA companies against Wintel and IBM and used employee scale to further underscore his point that Amazon's employee base is far outstripping Microsoft's.

What are the new S-curve candidates? There are 3 phases at the beginning of the S-curve, firstly getting the technology to work at all, secondly finding product/market fit and thirdly "pouring on rocket fuel." Benedict identified 4 technologies that are at different stages:

- Autonomy (which doesn't really work yet)
- Mixed Reality (which is just reaching the point where technology is starting to work but it's not really a product yet)
- Crypto-currencies (which have a proven, working technology but we're still working out where the product/market fit is)
- AI (which is well established in the "rocket fuel" phase)

Benedict went on to describe his view on what AI is and admitted that he feels calling it "AI" is unhelpful. We should instead be talking about Machine Learning, Automation and Enabling layers. He used the analogy of Databases to describe how moving from paper records to electronic record management and using structured queries such as "find me all the customers in a particular Zip code area" became an enabler for multibillion dollar businesses, ultimately leading to the Business Intelligence solutions of today. Today, "nobody looks at Starbucks and says, that's a database company." The database is an enabler in the same way that Machine Learning is a future enabler.

Future of AI - Measuring Progress and Preparing - Jim Spohrer

This year Ntegra hosted a full day programme of presentations at the Stanford Park Hotel. The schedule started with a thought-provoking presentation from Jim Spohrer, IBM Director of Cognitive Open Technology based at Almaden, San Jose, one of IBM's worldwide research labs. The centre opened in 1986 and continues the research started in San Jose more than fifty years ago. The phrase "Silicon Valley" was first seen in print around 1970 but the origins and heritage of the region stretch back to when Stanford University, NASA Ames and IBM Research in San Jose were doing pioneering work on silicon-based transistors.

As well as being Director of IBM Global University Programmes Worldwide, Jim has been the Director of Almaden Services Research and between 2000 and 2002 was the Chief Technology Officer for IBM Venture Capital Relations.

During the last 6 years Jim has been campaigning internally to get IBM to embrace open source in AI, latterly with great success. Half of his team are at IBM Watson West in San Francisco, working to transform IBM into a cognitive enterprise. Jim says that "AI is hard," especially for "old style" enterprises. As an example, JP Morgan Chase have just hired a new VP of Infrastructure from Google to transform their organisation into a cognitive enterprise. Google is more advanced, using AI internally just about everywhere (and they open-source much of their research). Facebook, Amazon and Microsoft are all pushing in the same direction with varying degrees of success. IBM is considered most mature in the B2B space.

With a technical background, Jim likes to pursue collaborative research with academia. IBM now has 15 global research laboratories with around 3000 researchers costing about $1Bn. They are currently making twice that from licence revenues! When asked about IBM's future considering recurring poor quarterly results, Jim was upbeat - referring to IBM as a "sleeping dragon" that's about to awaken, using its size and cumulative research to push ahead in many areas. This was not a view shared by the panel later in the day who rejected the assertion saying that IBM had 'had its day'. Since returning from the tour there have been an increasing number of 'good news' stories appearing in the press about IBM, so we wonder if perhaps there is more to come from IBM than general perception dictates?

AI, Deep Learning & Compute Power

Artificial Intelligence is popular again. You can often see it mentioned on billboards in SFO. However, pattern recognition does not equal AI. Deep learning only works if you have lots of data and compute power. We finally have lots of data and compute power so finally, deep learning for pattern recognition is working well. However, AI is more than deep learning for pattern recognition. AI requires common sense reasoning, which will take another 5-10 years of research to deliver. How do we know this?

- Look at the AI leader-boards [on Kaggle.com]
- Read 'The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World' by Pedro Domingos.
- Look at the free online cognitive classes at https://cognitiveclass.ai/

Jim purports that we should pose and address the following questions:

- What is the timeline for solving Artificial Intelligence and Intelligence Augmentation (IA)?
- Who are the leaders driving AI progress?
- What will the biggest benefits from AI be?
- What are the biggest risks associated with AI, and are they real?
- What are the implications for stakeholders?
- How should we prepare to get the benefits and avoid the risks?

Deep Learning for AI Pattern Recognition depends on massive amounts of 'labelled data' and computing power (available since ~2012). Labelled data is simply

input and output pairs (such as a sound and word, image and word, English sentence and French sentence, or road scene and car control settings). Labelled data means having both input and output data in massive quantities. For example, 100K images of skin, half with skin cancer and half without, are needed to learn to recognise the presence of skin cancer.

Thanks to Moore's Law, every 20 years compute costs are down by 1000x. This coupled with ML developments will stimulate and enable the growth of personal and digital assistants which will become commercially viable from about 2020 and in widespread use within 20 years. Some vertical applications may become mainstream much earlier than expected. For example, we are already seeing growth in voice enabled devices using Siri, Alexa and 'Hey Google' capabilities.

AI Leaders

Jim said, "Watching progress on open AI leader boards is like gazing into a crystal ball".

Cost of Digital Workers

- Moore's Law can be thought of as lowering costs by a facotr of a....
  - Thousand times lower in 20 years
  - Million times lower in 40 years
  - Billion times lower in 60 years

Smarter Tools (Terascale) - (T)

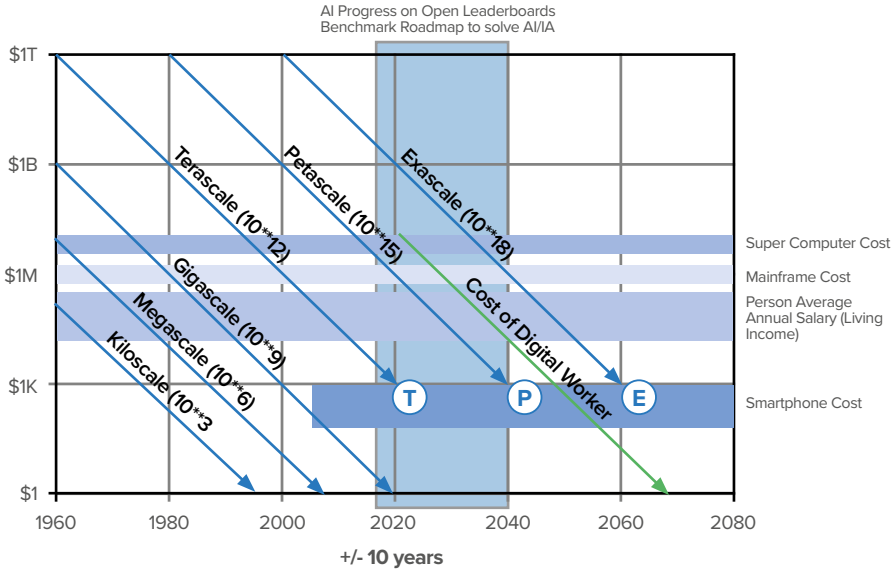
- Terascale (2017)= $3k
- Terascale (2017)= ~$1k

Narrow Workers (Petascale) - (P)

- Recognition (Fast)
- Petascale (2040) = ~1k

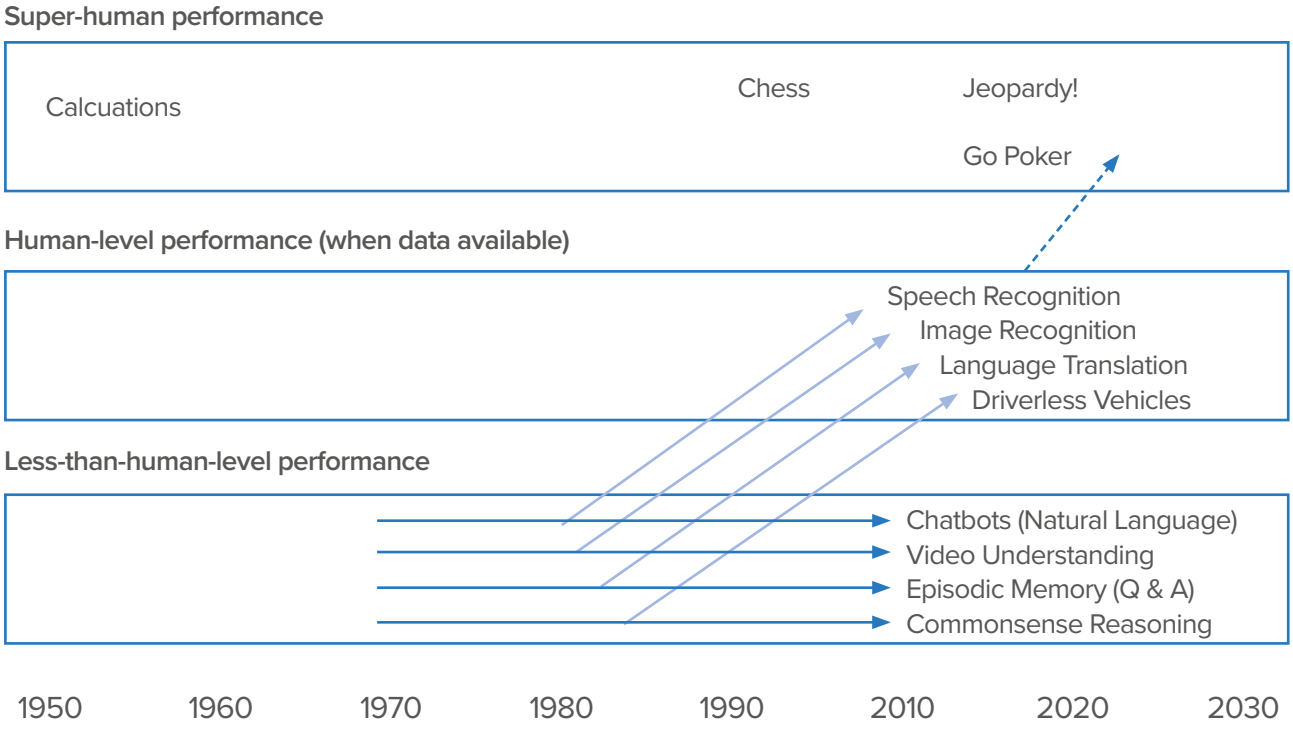
Broad Worker (Exascale) - (E)

- Reasoning (slow)
- Exascale (20260) = ~$1k



Who is winning? See: https://www.technologyreview.com/s/608112/who-is-winning-the-ai-race

Benefits & Implications



AI will undoubtedly facilitate easier access to expertise & better choices. More specifically:

- "Insanely great" labour productivity for trusted service providers
- Digital workers for healthcare, education, finance, etc.
- "Insanely great" collaborations with others on what matters most
- AI for IA = Augmented Intelligence and higher value co-creation interactions

This is opposed by the risks of job loss & the emergence of super-intelligence, potentially with the ability to go rogue. Shorter term risks include de-skilling of the work force and lower cost of certain attacks, for example spear fishing, allowing bad actors to automate tasks that were previously labour intensive.

In order to fully realise the reality of AI and make best use of it all stakeholders need to be involved - "The best way to predict the future is to inspire the next generation of students to build it better". It is essential that we consider everyone a stakeholder in AI due to its revolutionary nature. This engagement should be cross societal from individuals to families, small businesses to large multinationals, advisory groups to government and technical specialists.

Considering everything that Jim outlined, it is essential that the risks are mitigated. The report "The Malicious Use of Artificial Intelligence: Forecasting, Prevention and Mitigations" recommends:

- AI researchers should acknowledge how their work can be used maliciously
- Policy makers need to learn from technical experts about these threats
- AI world needs to learn from the security world how to best safeguard against threats
- Ethical frameworks need to be developed and followed
- More inclusive discussions across AI scientists, policy makers, ethicists, business and the general public

The Future

By 2036 there will be an accumulation of knowledge as well as a distribution of knowledge in service systems globally. As there is knowledge accumulation we need to ensure that service systems at all scales become more resilient. This will lead to the capability of rapid rebuilding of service systems across scales. Key to this are T-shaped people who understand how rapidl rebuilding works so that knowledge has been chunked, modularised and put into networks that support the rapid rebuilding.

To prepare to get the benefits and avoid the risks, this is what Jim tells his students, to provoke their thinking about the cognitive era:

1. 2015 - about 9 months to build a formative Q&A system - 40% accuracy;
   - another 1-2 years and a team of 10-20 can get it to 90% accuracy by reducing the scope
   - today's systems can only answer questions if the answers are already existing in the text explicitly
   - debater is an example of where we would like to get to in 5 years: https://www.youtube.com/watch?v=7g59PJxbGhY
   - more about the ambitions at http://cognitive-science.info
2. 2025: Watson will be able to rapidly ingest any textbooks and produce a Q&A system
   - the Q&A system will rival C-grade (average) student performance on questions
3. 2035 – as above, but rivals C-level (average) faculty performance on questions
4. 2035 - an exa-scale of compute power costs about $1000
   - an exa-scale is the equivalent compute of one person's brain power (at 20W power)
5. 2035 - nearly everyone has a cognitive mediator that knows them in many ways better than they know themselves
   - memory of all health information, memory of everyone you have ever interacted with, executive assistant, personal coach, process and memory aid, etc.
6. 2055 - nearly everyone has 100 cognitive assistants that "work for them"
   - better management of your cognitive assistant workforce is a course taught at university

In 2015 we were at the beginning of the beginning of the cognitive era. In 2025 we will be middle of beginning, easy to generate average student level performance on questions in a textbook. In 2035, we will be at the end of beginning (one brain power equivalent), easy to generate average faculty level performance on questions in textbook.

http://www.slideshare.net/spohrer/spohrer-ubi-learn-20151103-v2

By 2055, roughly 2x 20-year generations in the future, the cognitive era will be in full force.

Cell-phones will likely become body suits, with burst-mode super-strength and super-safety features:

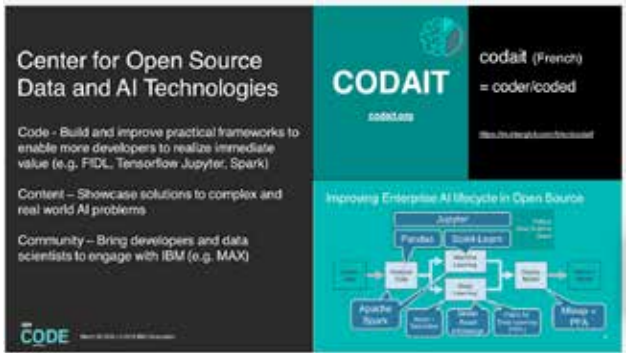
- Suits - body suit cell phones
- Cognitive Mediators will read everything for us and relate the information to us, and what we know and our goals.
- Combined personal coach, executive assistant, personal research team.

The key is knowing which problem to work on next. See this video for the answer (energy, water, food, wellness): https://www.youtube.com/watch?v=YY7f1t9y9a0&index=10&list=WL

Resources

To prepare for the future of AI, Jim recommended several resources:

- IBMs Centre for Open Source Data and AI Technologies (CODAIT.org)



- GitHub, "…isn't just for programmers. Everyone should have a GitHub account"

  (github.com) is like Wikipedia with code, that executes (code puts knowledge into action). In 10 years programming will be like chess, it will be solved!
- Kaggle, "…leader-boards for coding competitions, acquired by Google 2 years ago"

VC Panel

On Wednesday afternoon, we hosted our first Ntegra Greenside CIO Executive Panel, debating the critical issues that our delegates face and trends from the valley. The panel comprised seven executives from venture capital and start-ups, including Jeb Miller (from ICON Ventures), Jeff Pashalides (of Sequoia), Tom Rikert (from Next World Capital), Alex Fielding (CEO and Founder of Ripcord), Tim Eades (CEO of vArmour and Investor), Jay Srinivasan (CEO and Co-founder of Spoke) and JF Gauthier (CEO of Startup Genome).

JF Gauthier gave a brief description of Startup Genome, his 5th start-up, which promotes "Ecosystem and Scaleup Science" to provide help to companies around the World in building "their own Silicon Valley" for innovation and job creation. JF believes that the economic value of the "technology revolution" is currently highly concentrated and needs more distribution to enable the next phase of growth and scale. Despite increasing global competition for resources and markets, JF says that real economic value comes from larger scale-ups and bigger ecosystems that perform better.

The panel were asked: "What are the minimum constituents of a successful ecosystem?"

Discussion and responses covered ecosystem connections and networks, and the observation that businesses with greater than 50% foreign customers (globally focussed scaleups) outperform nationally focussed firms. The panel asserted that network connections are key to success regardless of ecosystem size: global knowledge enables creation of globally leading start-ups (through global connectivity). "How many people do you know and have a 'quality' relationship with?" In Silicon Valley connectivity and high-value relationships within the ecosystem are key to progression.

There is a mindset in Silicon Valley that encourages extremely high levels of ambition, enabling people to 'think the unthinkable' and where failure is not frowned upon but valued as good experience. This plays into the 'war for talent' discussion where there is an increasing desire to employ experienced software engineers over and above the best qualified Stanford graduates. The panel agreed that where talent was concerned you can't compete on price. Tim Eades reminded us that, due to difficulties getting visas, the market for developers is become increasingly "inward-focused" and there is a real danger that the talent pool may run dry, hence the race to, and cost of securing experienced developers.

Alex Fielding told us that Ripcord bootstrapped itself through seed to series 'A' spending around $600k. Just about everything was "begged, borrowed, or stolen" except where people were concerned. All the money before Series 'A' was spent on people, not "fancy things". Alex says there is a need to be creative in recruiting, saying wherever possible recruit your friends. "Real entrepreneurs go through walls or around them". Alex was connected to 3 heads of Silicon Valley Robotics clubs. He says recruiting them was effective but the additional intangible value they bring through their networks is priceless. "Always reward the stars, but the most painful thing is exiting the poor performers, the bottom 5%. Avoid politics and bureaucracy and avoid recruiting the 'first idiot'".

Getting back to the global distribution of funding, it was highlighted that most Venture Capital firms now have at least two global locations, if not more. For example, Tom discussed Next World Capital's global spread and commitment. The panel were all interested and keeping their eyes on the distortion in VC funding caused by large streams of Chinese money entering the system. They were also critical of some of the recent very large capital injections into very new businesses that had little more than an idea on a PowerPoint presentation. Jeb talked about "value capital efficiency", saying it is important to 'right size opportunities for start-ups to exit beneficially' in future. Overfunding has hurt the ecosystem previously.

"Venture capital is 'patient' capital" and VCs can drive its value by bringing the right context technologically to a business it has invested in. It is important not only have a vision, but to map the path to realising that vision. Tim also mentioned how important it is to understand the DNA of the organisations companies are selling to and where possible, sell to companies who are used to dealing with emerging tech and start-ups. Jeb used the example of Exabeam who built their 'betas' specifically for the companies they wanted to sell to. Tom also stressed that start-ups should target customers who can provide good public reference early in the relationship. The panel agreed that in some cases the references and visible traction are worth more to start-ups than the early revenue they can achieve.

Jay said that working with start-ups provides enterprises with “brilliant customer service” and they should be happy to find ways to match the pace, commitment and flexibility that start-ups have to maximise the benefit of working with them. It’s no good having an 18-month RFI/RFP procurement process if the start-ups they are working with only have a 12-month runway.

The panel also brought up the efforts that Government has been putting into making it easier to procure from early to market ventures. The Federal Other Transaction Authority (OTA) describes the streamlined procedures that federal agencies may use to procure innovative research or prototypes, without the constraints of a typical contract, grant or cooperative agreement. This flexibility has made OTA an increasingly popular choice for federal acquisitions in recent years. OTA helps open the door for contractors to partner with the Government in new and exciting areas. OTA allows for much greater speed, flexibility and accessibility in performing research and prototype projects.

In-Q-Tel (a specifically focused VC) identifies and partners with companies developing new technologies to help deliver solutions to the Central Intelligence Agency and the broader U.S. Intelligence Community to further their missions.

The Government has ‘a lot of data’ and it is keen to promote this. The cost and sourcing of good training data, datasets used in training Machine Learning neural networks can be prohibitive (“if you’re not a Google or Facebook”) so making large volumes of data available to start-ups through programmes like OTA is very ‘symbiotic’.

Jeff brought us back to the talent question saying, “AI is starting to gain real traction in enterprises, so data scientists are starting to be recruited into IT functions”. This led to a brief conversation about data scientists vs. data analysts, whether there is a natural progression from analyst to scientist and, therefore if not, will the next talent war be about data scientists? This resonated with Jim Spohrer’s comment that “Today’s AI ‘experiments’ are tomorrow’s Computer Science”, and we all need to be preparing for that future. Automating “more of the automation” is what AI can really drive. Alex repeated the assertion that current AI techniques and training already exceeds human capability in document understanding, at 60%.

Tim discussed his principles for investing in Cyber start-ups saying he avoids hardware, mobile, infrastructure, etc. and looks to invest in software only firms. He sees around fifty start-ups each year and invests in just 2 or 3. It is very important that the firms solve real business problems with real solutions, and the panel confirmed that Blockchain “doesn’t yet”. Tim likes to increase the value of domain experts in founding teams by hiring good people from target customers. He used Mark Weatherford, SVP & Chief Cybersecurity Strategist at vArmour as a great example.

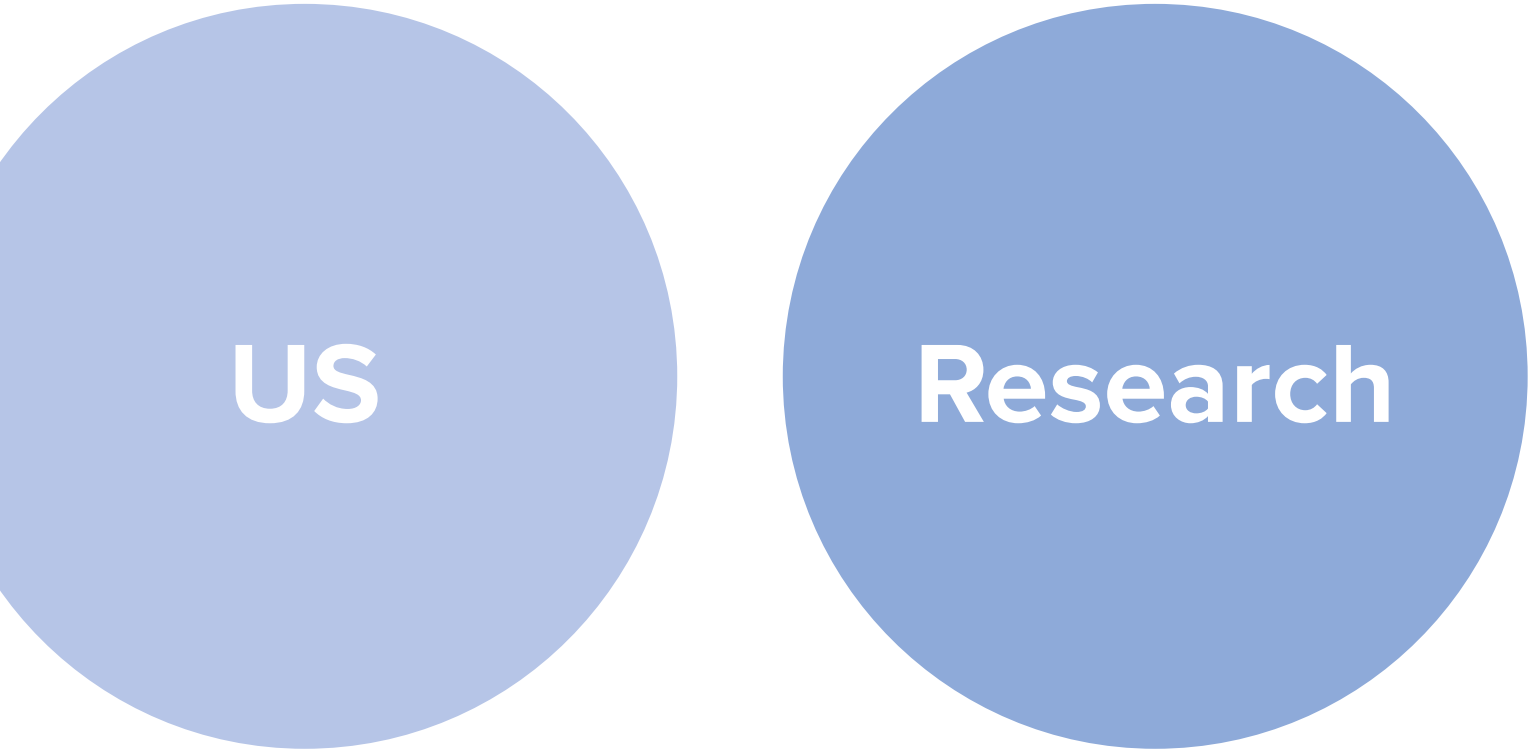
There was an open discussion about Open Source which to an extent is used by most if not all start-ups in the Valley. A lot depends on the licence conditions and whether consumers need to contribute back to the source or there is a cost to using software as part of a commercial offer. AWS provide a lot of ‘Open Source Services’ which are not fully open and not really commercial, making it difficult to navigate, especially for start-ups. The discussion led on to governance and the increasingly important role of Chief Data Officer responsible for legislative requirements.

Prompted by Andy Langley’s closing question, the panel went on to discuss the move from platform start-ups to vertical start-ups (e.g. Uber for mobility or Airbnb for hospitality) which resonated with Ntegra Greenside’s experience visiting Israel (on our forthcoming tour) which heavily verticalized. The conversation drew parallels with Benedict Evans’ Futures presentation.

Other interesting resources, references and insights that were referred to during the Panel discussion:

- Apache OpenNLP: https://opennlp.apache.org
- TrustArc: https://www.trustarc.com
- SpiceWorks: https://www.spiceworks.com
- Everything under cost control
- Open Source is changing the patch - think:
- Licence
- Licensing
- Licensibility
- Advise kids to:
    - be a product manager
    - work in AI
- Prioritise experience at low cost
- AI hype peak was 18 months ago
- Marketing through the crowd (community marketing model)

Monday 21 May



Tour

Ntegra Greenside

Knowledge for IT Leaders

Andreessen Horowitz

Founded	2009 Headquarters: Menlo Park, California
Founder(s)	Ben Horowitz, Marc Andreessen
Investment Funding	Debt, Early Stage Venture, Late Stage Venture, Private Equity, Seed
Website	www.a16z.com
Sector	Angel Investment, Finance, Venture Capital
Key Points	• Andreessen Horowitz is a Silicon Valley based venture capital firm with $4 billion under management • Invest from seed to growth • Phrase of the day: 'What will be the next S-curve'

Overview

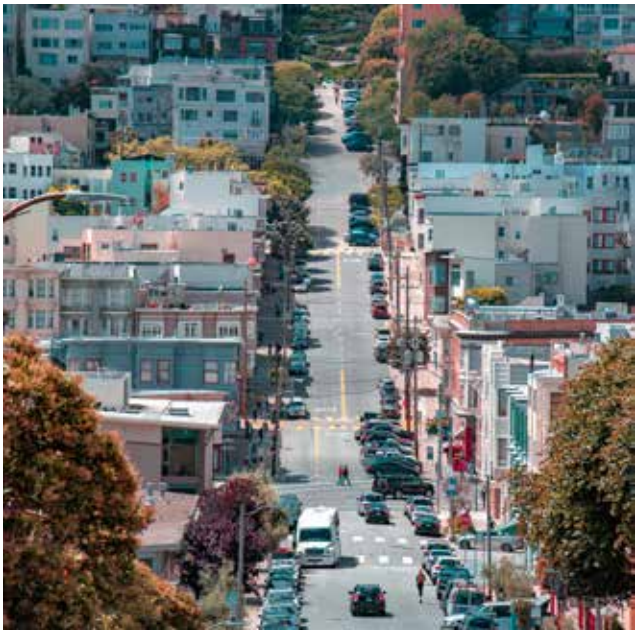
Andreessen Horowitz is a $4 billion venture capital firm that backs bold entrepreneurs, who move fast, think big and are committed to building the next major franchises in technology. Founded by Marc Andreessen and Ben Horowitz in 2009, the firm prefers to invest in technology sectors with a focus on enterprise software, back-end infrastructure, the Internet (infrastructure, consumer, business and mobile), cloud computing, social network, data-storage, consumer electronics and networking. Andreessen Horowitz assists the start-ups it invests in with everything from recruiting to public relations and is structured differently from most venture capital firms.

Technology

- Yubico
- Forward Networks
- GitHub
- Imply
- Mesosphere

Summary

Benedict Evans kicked off the tour with an overview of future trends that Andreessen Horowitz is tracking.



Yubico

Founded	2007 Headquarters: Palo Alto, California
Founder(s)	Stina Ehrensvard
Investment Funding	Total Funding: $30M Last Funding: Venture
Website	www.yubico.com
Sector	Cloud Security, Cyber Security, Hardware, Open Source, Software, Wireless
Key Points	• Offers USB and wireless authentication solutions • Simple, secure login, preventing unauthorised system access • Supports multiple authentication and encryption protocols on all devices and platforms

Overview

Yubico provides simple and secure login, preventing unauthorised access to computers, servers and internet accounts. Supporting multiple authentication and encryption protocols on all devices and platforms, YubiKeys protect access to user accounts for the world's largest enterprises with a simple touch, with no driver or client software needed. Yubico is a leading contributor to the FIDO Universal 2nd Factor open authentication standard. Yubico's technology is used in more than 150 countries.

Technology

USB YubiKey works natively with PC's and NFC enabled Smartphones NFC (currently Android only).

YubiKey for macOS (10.12 and Later) enables secure login to Macs using the smart card (PIV) mode. For Mac OS X (10.9 and Later) YubiKeys can be configured by setting up Challenge-Response using the Yubico Pluggable Authentication Module (PAM).

Applications

Works across a range of scenarios; Computer Login, Remote Access & VPN, Identity Access Management, Password Management, Privileged Access, Online Services, Developer Tools and Encryption.

Case study from Yubico:In 2009 Google was the target of sophisticated cyberattacks capable of circumventing traditional security controls. With a lack of viable two-factor authentication (2FA) options to effectively prevent these attacks and account takeovers, Google began working closely with Yubico to extend the capabilities of the YubiKey two factor authentication technology to also include public key cryptography. Through this collaboration, Yubico and Google co-created a strong authentication protocol based on the concept of a single un-phish-able key to secure all services. It is this work that later became an open standard adopted by the FIDO alliance called the FIDO Universal 2nd Factor (U2F) standard.

One single YubiKey can secure a multitude of online services with no user information or private keys shared between the service providers. There is no reliance or requirement for mobile connectivity, cellular devices, mobile apps or manual code entry.

Observations

Delegates were concerned at the additional overhead required to provision and manage physical keys, and it was felt that questions regarding lost keys were not adequately responded to. There was a consensus that the solution would work well for a smaller business or individual team.

All Ntegra Greenside delegates were presented with free Yubikey Neo and 4C Nano keys to enable offline trials, and we would be most grateful if you could provide any feedback on your experience using them?

Forward Networks

Founded	2013 Headquarters: Palo Alto, California
Founder(s)	Brandon Heller, David Erickson, Nikhil Handigol, Peyman Kazemian
Investment Funding	Total Funding: $27.1M Last Funding: Series B
Website	www.forwardnetworks.com
Sector	Information Technology, Software
Key Points	• Delivers a mathematically accurate network copy to simplify network operations and provide network assurance • Provides network visibility, policy verification, and change modelling • Next generation network management aimed at complex modern highly distributed networks

Overview

Forward Networks delivers an innovative platform that provides network visibility, policy verification and change modelling. Designed to help network teams eliminate costly network outages, the Forward Networks' platform enables engineers and operators to easily visualise and search complex networks, quickly debug problems, verify network-wide policy accuracy and predict network behaviour prior to making changes to production equipment. Fortune 100 financial institutions, telecommunications and IT service providers use Forward Networks to 'de-risk' business operations and drive operational efficiency.

Technology

Forward Networks' technology includes Search, Predict and Verify capabilities to manage networks and anticipate issues (Intent-Based Networking, a term coined by Forward Networks).

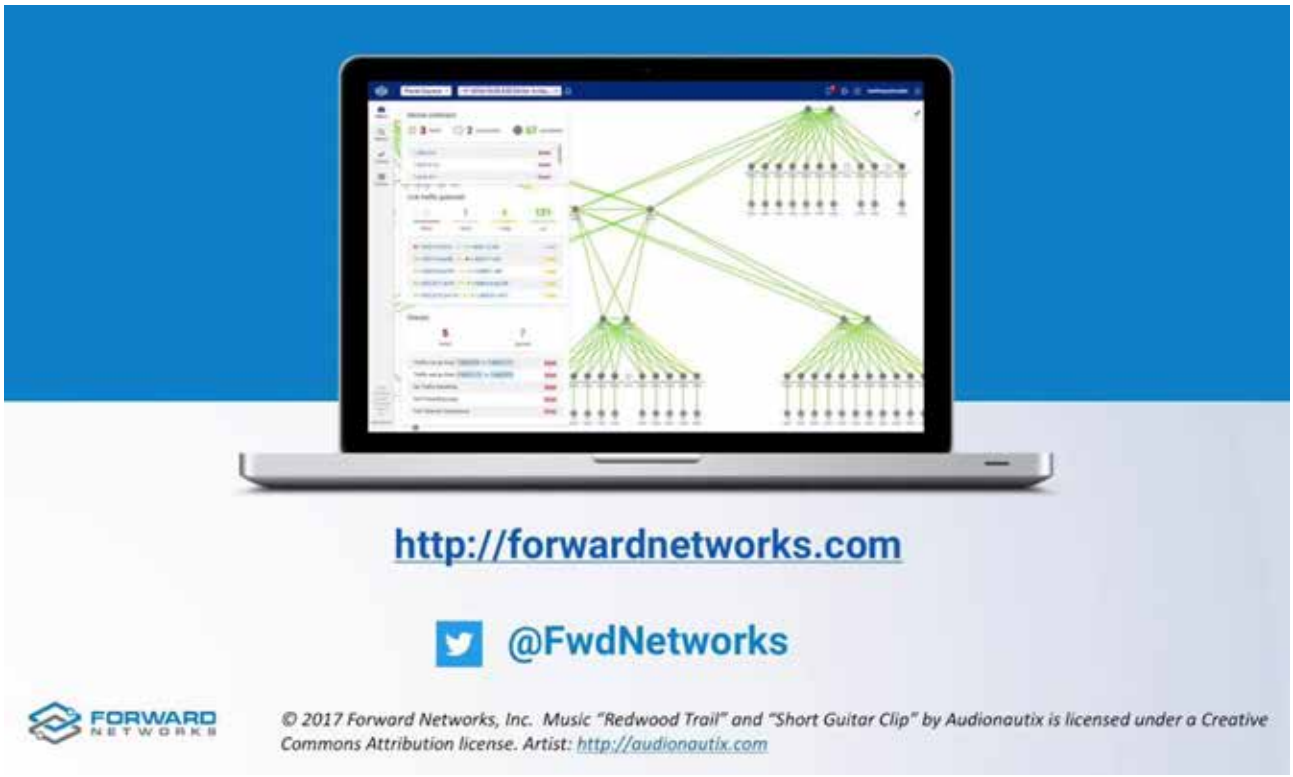
The solution can be deployed on-premise as a virtual machine (VM), or in Cloud to provide a 'single point of truth' about the network, enabling administrators to ask questions such as, "can particular servers only be reached using HTTP(S) protocols?", and allowing them to run a mathematical verification of network configurations under all conditions. The solution also enables sandboxing and model forking to assess whether changes may impact network quality of service or introduce lateral contagion in the event of a security breach.

Applications

Aimed at managing complex distributed networks and preventing issues by using model-based verification of network performance. Based on three core components, Forward, Search and Verify, Forward Networks reduces operational overheads and improves efficacy of change management. The platform enables hidden errors and incorrect device configurations to be identified easily.

Observations

The networking industry is at a critical point with network administrators and operations teams struggling to manage much larger and more distributed networks. Traditional network management solutions and network change management software is no longer able to tackle the complexity at scale or speed of business, leaving the potential for major network outages and business disruption.



GitHub

Founded	2008 Headquarters: San Francisco, California
Founder(s)	Chris Wanstrath, PJ Hyett, Tom Preston-Werner
Investment Funding	Total Funding: $350M Last Funding: Non-equity Assistance 2 acquisitions
Website	github.com
Sector	Cloud Computing, Developer Tools, Internet, Project Management, Software
Key Points	• Web-based Git repository hosting service • Distributed revision control and source code management • Functionality of Git

Overview

GitHub is a web-based Git version control repository and Internet hosting service that offers all the distributed version control and source code management functionality of Git, as well as adding its own features. It provides access control and several collaboration features such as bug tracking, feature requests, task management and Wikis for every project. Its enterprise version hosts software repositories behind corporate firewalls and its customers include NASA and SAP.

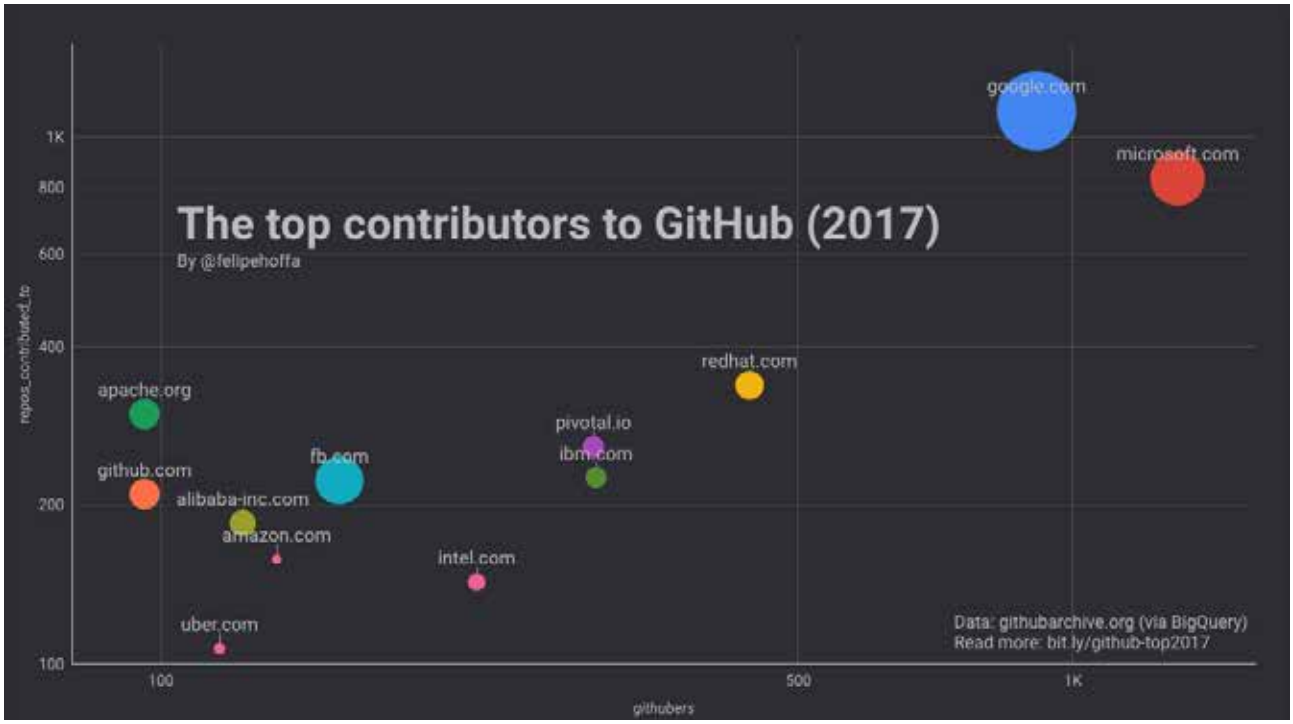
A16Z had previously presented GitHub in 2017 but wanted to provide more insight to the value and potential impact of meta-data within the platform (about developer trends, how companies run internal development projects as if they were open-source and where organisations are writing software that has already been developed). At the time of the presentation we were unaware that within a fortnight, this rich and valuable seam of information would be acquired by Microsoft for ~$7.5bn.

Technology

GitHub is a software development platform comprising 68 million repositories and 28 million users with 1.6 Million organisations. They retain over 500 terabytes of software development data accrued over 10 years. The patterns yielding insights that emerge from this volume of data is enabling GitHub to transform from a Software Company into and Innovation Company.

Applications

Originally GitHub was a place where developers could work privately but eventually it spread to the open source community. GitHub began exclusively as open source, but soon many of its customers began requesting the capability to work on projects privately. GitHub then created GitHub teams and in March 2017 they launched Business Hosted. Business Hosted (the enterprise equivalent of GitHub Teams) enables enterprises to securely extend GitHub into their organisation, enabling them to take advantage of all the benefits of using GitHub but within a secure environment to which the enterprise can control access.



Over time some organisations wanted to host GitHub within their internal IT Infrastructure behind their own firewalls so Business On-Prem was developed and has just been released as a hosted virtual appliance, on premise.

In March 2017 the GitPrime app was released. It is an app that tells the user who is doing work and to what degree. It analyses data from the GitHub database and code repositories and measures team and individual productivity, specialisms, code quality, individual developer work patterns and over time much more. Other actionable data include:

- Most valuable code
- Most volatile code
- Most reused functions
- Most popular languages
- Most edited files
- Most used tools
- Best collaborators
- Suggested reviews

Observations

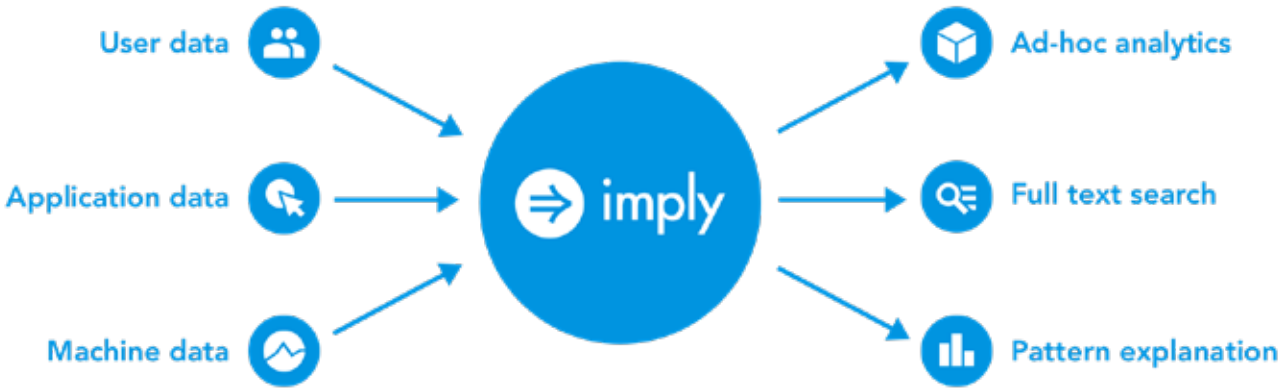
Open Source development practices are becoming a hallmark of modern, competitive engineering organisations. GitHub has been a keystone supporting Open Source development communities for more than a decade. Their platform and data provide their ‘unfair advantage’ and USP. Microsoft has said that it is committed to continued investment in the GitHub roadmap and that it is buying GitHub because they ‘like it’. It will be interesting to see whether they are able to retain users and stick to their promises. Before the acquisition announcement, Jim Spohrer (IBM) suggested that everyone should become a GitHub user, saying that >50K IBMers have accounts. We wonder how this may play-out in future?

Imply

Founded	2015 Headquarters: Millbrae, California
Founder(s)	Fangjin Yang, Gian Merlino, Vadim Ogievetsky
Investment Funding	Total Funding: $15.3M
Website	imply.io
Sector	Analytics, Data Visualization, Open Source
Key Points	• Operations centre platform • Enables collection, analysis and diagnosis of business operations in real-time

Overview

Imply is an operations centre platform that allows users to collect, analyse and diagnose business's operations in real-time. Imply is commonly used for customer behaviour, security operations, system performance, network traffic, manufacturing, digital media and network data. Imply enables anyone in an organisation to create and view reports at any level of detail on real-time and historical data, providing complex aggregations, groupings and filters with no experience required.



Technology

Imply is a high-performance analytics solution for time series & telemetry data. It is used to consolidate, analyse and explain critical operational data, providing operational analytics across structured, un-structured and semi-structured data. Its core is based on Druid, a widely adopted technology that uses Lamda functions created by Imply's founders. Imply is an Enterprise ready wrapper around the Druid core which can be run on-prem or in the cloud, combining a powerful data engine with intuitive user interface.

Applications

Traditional solutions primarily focus on monitoring or search. Imply adds analytics and intelligence:

- Ad-hoc grouping and filtering, and fast numerical aggregation
- One-click explanation of trends and anomalies

Imply can handle Realtime and historical data in many different formats, supporting differing optimisations across structures, semi- and un-structured data such as logs, metrics and events.

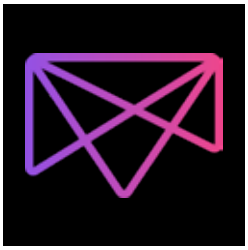
Observations

Imply drives business opportunity through provision of unified analytics, search and intelligence. It is designed to be used by anyone in an organisation and is optimised for timeseries and telemetry data, supporting network security and traffic analysis, user behaviour and industrial/IoT analytics.

Imply presented a compelling case study describing how PayPal improved their merchant behavioural analytics by consolidating data and providing a unified access point for users. Real-time data about all merchants, products and sales were streamed into Imply as well as historical batch data. Analysts were able to generate live ad-hoc reports.

Product managers were able to A/B test product changes and marketing could immediately measure campaign impact. End-users could also self-service their own questions and as a further value-add, PayPal is now providing an analytics product powered by Imply for their merchants to use.

Mesosphere



Founded	2013 Headquarters: San Francisco, California
Founder(s)	Benjamin Hindman, Florian Leibert, Tobi Knaup
Investment Funding	Total Funding: $247.3M Last Funding: Series D $125M on 7/5/2018 2 Acquisitions
Website	mesosphere.com
Sector	Cloud Computing, Cloud Infrastructure, Data Centre Automation, Enterprise Software
Key Points	• Cloud operating system that treats the data centre as "one big computer" • Spans all servers in a physical or cloud-based data centre • Runs on top of any Linux distribution

Overview

Mesosphere abstracts the entire datacentre into a single pool of compute resources, simplifying orchestration of distributed systems at scale. Mesosphere's Datacentre Operating System (DC/OS) is based on the open-source Apache Mesos distributed systems kernel providing a common substrate for cluster computing and fine-grained resource sharing in the data centre. Mesos is the foundation for several distributed system frameworks used by Twitter, Airbnb and eBay.

A16Z had previously presented Mesosphere to us in 2016 but felt it valuable to revisit the company in light of the significant advances that they have made during the intervening 2 years. Mesosphere is now able to unify public and private clouds into a single pool of resources and enables 'one-click deployment' and scaling of containerised applications and data services anywhere, instantly. DC/OS 1.11 now provides Kubernetes & hybrid support, making it the "abstraction layer of abstractions".

Technology

Apache Mesos is an open-source cluster manager that was originally developed at the University of California Berkeley RAD Laboratory. Mesos provides a scheduling mechanism that "offers" resources to frameworks for applications that run on top of Mesos. Each framework determines which resources to accept and what applications to run on them. This arbitration method allows near-optimal data locality when sharing a cluster of compute nodes amongst diverse frameworks. The Mesosphere DC/OS spans all the servers in a physical or cloud-based datacentre and runs on top of any Linux distribution. DC/OS provides container automation and governance for both 'stateful' and state-less distributed workloads including Docker containers, Big Data and traditional applications, in a single framework.

By having Mesos handle low-level infrastructure logic and delegating the application-specific logic to the framework Mesos can easily scale to tens of thousands of nodes, which is why it has been reliably powering some of the largest web-scale platforms in the last couple of years.

As new distributed technologies are introduced to the market organizations can easily introduce and adopt new technologies as applications or frameworks (such as Kubernetes), and include any application-specific scheduling and operational logic, making Mesos a true future-proof platform for distributed technologies.

Applications

DC/OS is an enterprise class hybrid-cloud operating system for elastic applications that ensures high datacentre utilisation and makes executing and managing applications in complex environments as easy as launching applications on a laptop. It addressed the complexity of hyper-scale operations and supports developer and data agility.

DC/OS has a powerful Command Line Interface (CLI) and Graphical User Interface (GUI) to enable data centre operation and a package manager to run containers and big data workloads in production.

-
- Mesosphere provides broad workload coverage enabling customers to run today's and tomorrow's applications including traditional J2EE, containers, analytics & ML
- Application-aware automation ensures customers can automate workload-specific operating procedures to "as-a-Service" anything from Kubernetes to data services
- Intelligent resource pooling lets customers optimise workload density for highest utilisation with resource guarantees
- Unified hybrid cloud operations enable customers to securely manage cloud, data centre and edge infrastructures from a single control plane

Whether it's traditional or modern data-rich applications made with containers and fast data services, DC/OS runs it all 'as-a-Service' for faster rollout of new projects like personalisation, machine learning and IoT.

Observations

Mesosphere's easy to install, run anywhere orchestration supports 'stateful' and stateless implementations with unlimited scaling potential. For customers with large investment in on-prem infrastructure, the ability to "burst to cloud" for peak workloads should enable a soft transition to cloud while depreciating their legacy assets.

ThoughtSpot

Founded	2012 Headquarters: Palo Alto, California
Founder(s)	Abhishek Rai, Ajeet Singh, Amit Prakash, Priyendra Deshwal, Sanjay Agrawal, Shashank Gupta, Vijay Ganesan
Investment Funding	Total Funding: $295.7M Last Funding: Series D
Website	www.thoughtspot.com
Sector	Analytics, Big Data, Business Intelligence, Enterprise Software
Key Points	• AI-driven Analytics for Humans • Relational Search Engine designed for data analytics • The days of waiting for reports and dashboards are over

Overview

A new breed of Relational Search Engine designed for AI-driven data analytics, for humans. ThoughtSpot enables the use of search to easily analyse data or automatically get trusted insights pushed “with a single click”.

ThoughtSpot is disrupting the BI industry with search-driven analytics built for the enterprise. Anyone can use ThoughtSpot with zero training to ask questions, analyse company data and build reports and dashboards, all in seconds. ThoughtSpot can combine data from on-premise, desktop, cloud sources and Hadoop, can scale to billions of rows and can be deployed in hours.

As well as presenting its own technology, ThoughtSpot hosted 2 other Ntegra curated companies, Diffbot and IdeaScale.

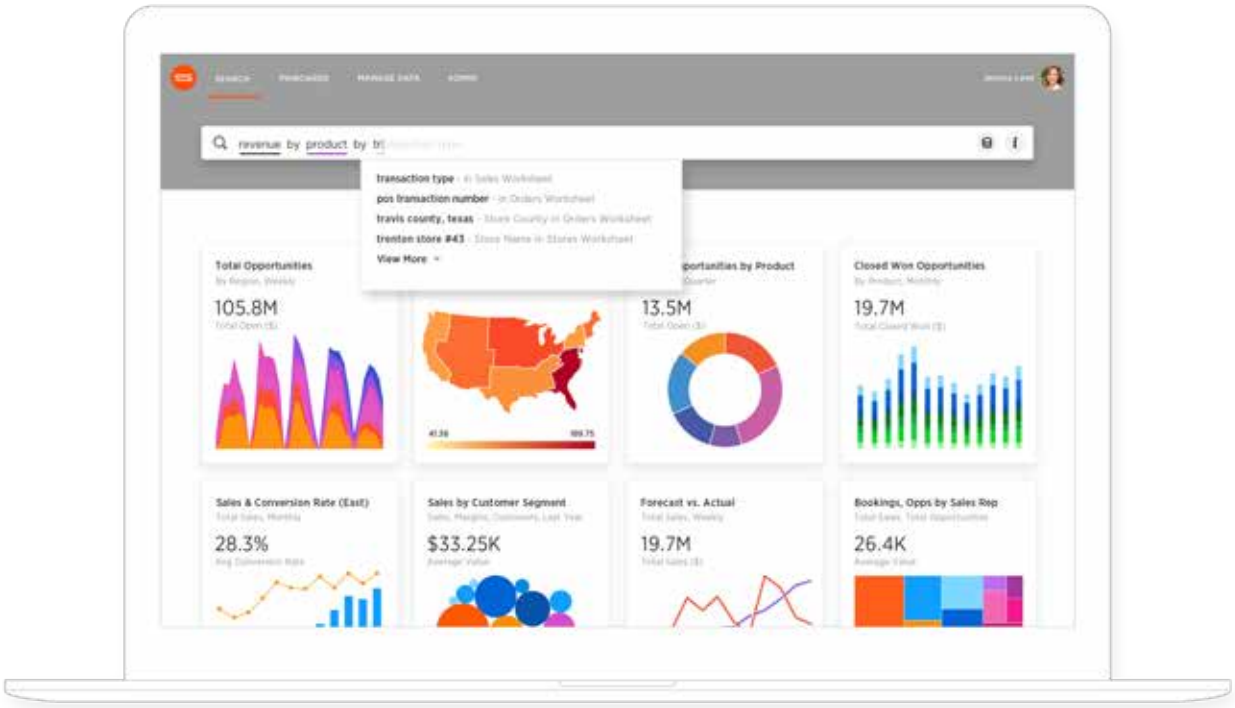
Technology

ThoughtSpot makes it easy for anyone to get answers from their company data in seconds. Simply search to analyse data and get automated insights with a single click. ThoughtSpot is designed for true self-service for everyone to answer both known and unknown questions in a secure and scalable environment.

ThoughtSpot combines a new relational search engine for data analytics with a custom-built in-memory relational data cache to provide sub-second response times to search queries run over billions of rows of data. It can be run as software in the cloud on an instance of AWS, on-prem as a horizontally scalable appliance solution or as software on a VMware-powered private cloud.

ThoughtSpot’s appliance is a stackable, 2U chassis appliance comprised of between 1 and 4 individual nodes that form a cluster. Appliances come standard with two Intel® Xeon® CPU E5-2690 v2 processors, two 10GbE network ports and two 1GbE network ports. To get started, simply install the appliance in a rack in a data center and connect to a dedicated LAN for system management, power on the appliance and begin loading data.

Building a relational search engine for numbers required a fundamentally different approach to enterprise analytics. While searching to analyse in ThoughtSpot is easy for users, there’s a lot of sophistication under the covers that makes it fast and accurate.



Relational Search Engine (guided search experience designed for fast analytics at scale):

- Guided Suggestions
- Relevant search suggestions based on each word you type, powered by your data, metadata and usage, all presented in real time
- Data Rank
  A ranking algorithm for search suggestions that leverages collective intelligence from prior usage, security rules and characteristics of data such as cardinality and data type
- Auto-Joins
  Understands your data model and joins tables with billions of rows on-the-fly to calculate answers across siloed data sources
- Synonyms
  Real-time keyword validation for customizable synonyms, homonyms and misspelled words for enterprise-wide usability
- Search Inspector
  Explains underlying calculations so you can easily verify and trust answers
- Advanced Analytical Search
  Use search to incorporate formulas, group sums, nested searches, many-to-many joins and multiple join paths

Visualization Engine (automatically chooses the best chart based on your search and presents them on live dashboards):

- Pop-Charts
  Intelligently chooses from dozens of visualisations as you type and presents them in real time
- Pinboards
  A drag-and-drop story building experience to help you organise insights on live dynamic dashboards
- Drill Anywhere
  Ad-hoc drill down and across any data dimension without the need for a hierarchy or pre-defined drill path

Collaboration (real-time comments on analyses and one-click sharing of charts and dashboards):

- Group & User Sharing
  One-click sharing with groups and individual users combined with an activity feed showing you real-time updates

- Presentation Mode
  An interactive view to present charts and pinboards that utilises every inch of screen's real estate
- Comments
  Chart annotations and pinboard messaging for collaboration in context
- Alerts and Notifications|
  Real-time alerts and scheduled notifications so you'll never miss the latest analysis

In-memory Calculation Engine (data cache optimised for search, query processing and calculations at scale):

- Query Optimisations
  Caches frequent searches for maximised query performance with zero tuning
- Parallel Computing
  Executes queries and performs joins and aggregations in parallel for answers in milliseconds
- Data Shard-ing & Replication
  Intelligent distribution of data based on database schema and table size to minimize join time while maintaining high availability

Distributed Cluster Manager (Highly reliable and scalable clusters complete with monitoring and real-time alerts):

- Web-Scale Architecture
  Fully distributed in-memory database on top of an infinitely stackable appliance built for hundreds of terabytes of data
- High Availability
  Data replication and automatic node failover for zero downtime
- Disaster Recovery
  Automatic backups and snapshots to restore your cluster
- Cluster Monitoring & Usage Analytics
  Real-time monitoring alerts and out-of-the-box usage dashboards for simplified user management and administration

Data Integration & APIs (Database connectors certified by the world's most popular ETL platforms and flat file uploading for fast parallel data loading):

- Database Connectors
  Native ODBC and JDBC connectors for use with your database and ETL platform of choice, certified by popular integration tools like Informatica Power Center and Microsoft SSIS
- ThoughtSpot Data Connect
  Instantly combine favourite cloud data sources with your on-premise data in ThoughtSpot
- Flat file upload
  Drag and drop Excel spreadsheets, CSVs, or any other delimited file for ad-hoc data mashup and analysis
- Fast Parallel Loading
  Bulk data loader capable of loading multiple terabytes of data per hour
- Embedded Analytics
  iFrame embedding of pinboards, charts and data directly into any portal or app

Enterprise-Grade Security (Secure and granular access control and a secure single sign-on experience):

- Row & Column Security
  Restrict data access down to the cell level based on customisable permissions
- Group & User Security

Comprehensive user and group governance that persists across charts and pinboards:

- LDAP Integration
  Native LDAP support for directory services software, including certification with Active Directory
- SAML Integration
  Easy integration with existing enterprise single-sign on software Marketing

In addition, SpotIQ leverages the power of ThoughtSpot's in-memory calculation engine to run thousands of queries on billions of rows of data and execute dozens of insight-detection algorithms. The most actionable insights for business are uncovered in seconds. It can auto-detect anomalies and outliers, identify relationships between measures that you didn't know about and discover trends on noisy data.

Applications

ThoughtSpot is a new breed of search engine designed for data analytics. As you type ThoughtSpot instantly calculates answers and presents the best visualisation for your search. ThoughtSpot's search engine uses collective intelligence to guide users to the right answers and is powered by a sophisticated calculation engine capable of performing aggregations over terabytes of data with sub-second latency. Complete with robust integration, fine-grained security rules and cluster management for reliability and massive scalability, ThoughtSpot brings instant answers to every business professional.

Observations

ThoughtSpot say they can connect with any on-premise, cloud, big data or desktop data source and deploy 85 percent faster than legacy technologies. BI & Analytics teams have used it to cut reporting backlogs by more than 90 percent and make more than 3 million decisions. Customers include Amway, Bed Bath and Beyond, Capital One, Celebrity Cruises, Chevron Federal Credit Union, De Beers, Insurethebox and Scotiabank. ThoughtSpot's mission is to enable analytics at "human scale" and put search-driven analytics in the hands of 20M users by 2020.

Gartner predicts that "augmented analytics," which includes using natural-language queries and visual-based data discovery, will be a key driver of enterprise purchasing decisions.

Platforms that slice and dice corporate data are not new, but demand for high-powered, easy-to-use analytics platforms has risen as the need for data analytics skills grows outside the IT department. Teams in marketing, finance, HR and other functions increasingly are expected to crunch data ranging from sales figures to weather patterns to help drive their decision making.

Other companies presented while at Thoughspot

- Diffbot
- Ideascale

DiffBot

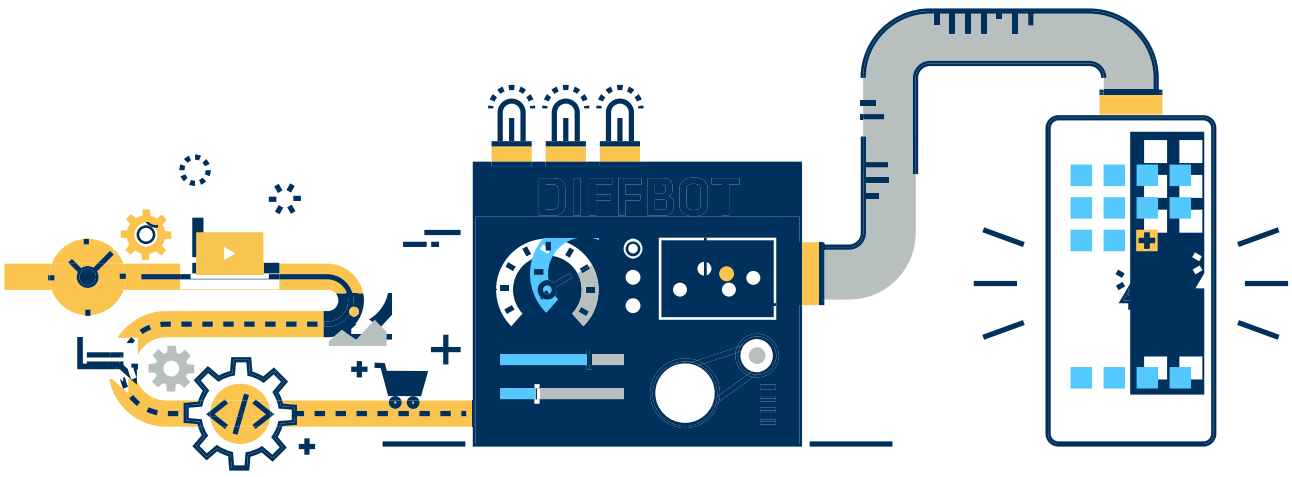
Founded	2011 Headquarters: Mountain View, California
Founder(s)	Michael Tung
Investment Funding	Total Funding: $13M Last Funding: Series A
Website	www.diffbot.com
Sector	Analytics, Developer APIs, Enterprise Software, Internet, Machine Learning, Mobile, Natural Language Processing, Semantic Search
Key Points	• AI start-up that provides knowledge as a service to power intelligent applications • Turns websites into data in seconds • Creating the world's largest database of people and companies

Overview

Using AI, ML & NLP, Diffbot provides software developers with the tools to extract and understand objects from any web page. They provide automatic and custom APIs + 'crawlbot' & bulk processing capabilities. Diffbot is a team of AI engineers building a universal database of structured information, to provide knowledge as a service to all intelligent applications. Whether you are building an app that uses web content, an enterprise business application or a smart robotic assistant, Diffbot can provide the data you need.

Diffbot was the first company to be accelerated in the StartX program that would become Stanford's premiere source for getting graduates 'to exit velocity' with their business.

Diffbot, which raised its first seed round in 2012, has set itself the goal of being the "Intel of data" for independent artificial intelligence application developers. With a new $10 million commitment led by Tencent, one of China's largest Internet companies, Diffbot has come a long way from solving the math problems that would form the core of the company's artificial intelligence software.



Technology

Diffbot's APIs automatically extract content from supported page types: articles, products, discussions, images and more. Diffbot uses advanced AI technology to retrieve clean, structured data without the need for manual rules or site-specific training. Crawlbot lets you apply any Diffbot API to an entire site, structuring hundreds or thousands of pages into a single searchable index. With bulk processing, extract structured data from hundreds to millions of URLs in a single job. You can also extract any data from any web page using easy-to-create custom rules and an instant API.

Applications

Companies like Google, Facebook, and Baidu, which are all working on artificial intelligence, have the benefit of massive amounts of data at their fingertips that they and their data entry employees can use to categorise and define the web in a language that AI software can later feed into their algorithms. Small companies who don't have the benefit of that data can turn to Diffbot. Unlike other artificial intelligence deep learning projects that rely on humans to classify web pages, Diffbot uses only the proprietary algorithms that it created itself and has refined over the years. The problem with the data that would be intelligences learn from and process, is that it needs to be structured in a way that the systems can recognise, and that's exactly what Diffbot does, "taking the Internet and converting it into semantic knowledge".

Observations

Diffbot is focused exclusively on getting better web data. Working without rules or training, they say there's no better way to extract data from web pages. By 2015 Diffbot's AI software was identifying data on the web with a 90% to 95% reliability so the company started spidering the web to speed up its data collection. Their goal is to get to trillions of discrete data points to provide a structured taxonomy for the entire internet. It's taxonomy already contains more than 1.5 billion objects and is adding 10 million objects per day. By comparison, Google's Knowledge Graph only recently passed 1 billion objects.

IdeaScale

Founded	2009 Headquarters: San Francisco, California
Founder(s)	Robert M Hoehn, Vivek Bhaskaran
Investment Funding	2 Acquisitions
Website	ideascale.com
Sector	Crowdsourcing, Innovation Management
Key Points	• Cloud-based innovation software platform • More than 25,000 customers and 4 million users • Software allows organizations to involve the opinions of public and private communities

Overview

In 2009 IdeaScale founders Vivek Bhaskaran and Rob Hoehn launched IdeaScale in tandem with President Barack Obama's Open Government Initiative. In its first year IdeaScale was adopted by 23 federal agencies. It served many organisations, including the Executive Office of the President of the United States. During the following year, the platform's adoption rate expanded to include more than 36 agencies as well as numerous private Enterprise companies. Government clients cited IdeaScale's "level of engagement as well as the platform's affordability" as one of the reasons that it was adopted early in the open government initiative. IdeaScale's current clients include EA Sports, NBC, NASA, Xerox and the White House.

Technology

IdeaScale's mission is to provide superior open innovation software (as a Service) to transform every business' products, processes or structure. IdeaScale is the largest cloud-based innovation software platform in the world with more than 25,000 customers and 4.5 million users. The software allows organisations to involve the opinions of public and private communities by collecting their ideas and giving users a platform to vote. The ideas are then evaluated, routed and implemented, making IdeaScale an engine of crowd-powered innovation.

Applications

IdeaScale allows organisations of all kinds to customise their innovation community to meet their own requirements. Communities are configured pre-launch to get the highest volume, best quality and most relevant ideas aligned to goals. Community members collaborate with one another as they suggest ideas while voting and commenting on the ideas of others. Integrations lower participation boundaries and make the experience more natural for all participants. The crowd provides ideas but you need a team to build proposals and prototypes and eventually deliver innovative products and services.

IdeaScale allows teams to form around an idea and for ideas to be assigned to the leader most able to research and manage an idea proposal. The crowd provides numerous valuable ideas but knowing which one to invest in is often a challenge. With IdeaScale's back-end functionality, administrators can

evaluate benefits, estimate potential costs, forecast future success with greater ease and deliver results with reliability. Delivering on the selected ideas is the best way to build a sustainable innovation program. IdeaScale integrates with numerous in-house project management and delivery modules to track & socialise the progress of any idea.

Observations

IdeaScale champions moving from a closed R&D style model to a collaborative, open model of innovation. Nearly 38% of IdeaScale customers have a dedicated innovation group in their organisation whose responsibilities span numerous purviews – most often research and developing ideas through intrapreneurship or acquisition, stewarding ideas through prototyping and testing, training their organisation in innovation best practices, distributing knowledge and aligning efforts. This means they're building systems that not only capture every idea no matter where it's come from, but you can prove and articulate its value across the company and increase its impact.

Robert used Elon Musk phrase "moats are lame" as an exemplar of open innovation thinking. The Tesla CEO was 'calling out' Warren Buffett, who uses "moat" to describe barriers to imitation that stave off competition. "If your only defence against invading armies is a moat, you will not last long," Musk continued. "What matters is the pace of innovation, that is the fundamental determinant of competitiveness." See: http://bit.ly/moatslame

Tuesday 22 May

US

Research

Tour

Ntegra Greenside

Knowledge for IT Leaders

Ripcord

Founded	2009 Headquarters: San Francisco, California
Founder(s)	Alex Fielding, Kevin Hall, Kim Lembo
Investment Funding	Total Funding: $49.5M Last Funding: Series B
Website	www.ripcord.com
Sector	Digital Media, Office Administration, Software
Key Points	• Robotic digitisation company that builds robots to digitise records and make them accessible anywhere, anytime • Uses sophisticated automation and software to provide customers with a secure, fast and all-inclusive digitisation solution at low, predictable prices

Overview

Ripcord claims to be the world's first robotic digitisation company using sophisticated automation and software to provide customers with a secure, fast and all-inclusive records management solution at low predictable prices. Ripcord was established in 2015 prompted by a customer problem in dealing with massive volumes of paper records. They are the first company to apply vision robotics to paper digitisation. Ripcord were incubated and accelerated through a NASA start-up mentoring programme at Ames Research Facility. They are building an East Coast operation with security hardened facilities for government agency staff and their roadmap puts them in the EU some time next year with APAC capabilities following on shortly after.

As well as presenting its own technology and providing a guided tour and demonstration of their facilities and scanning robots, Ripcord also hosted 2 other Ntegra curated companies, CachePhysics and Element Analytics.

Technology

Upon receipt, packaged physical documents are logged, assigned a unique barcode and tracked as it moves throughout the digitisation process. Ripcord organises and prepares documents for digitisation robots known as 'Workcells'. The Workcells quickly digitise each image in full colour at high resolution and produce a fully searchable PDF of each page. Records are securely stored and managed via Ripcord Canopy, an in house developed records management solution.

Ripcord's digitisation is divided into two subprocesses: the conversion process and the management process. Ripcord automates 80% of the conversion process, which includes paper handling, fastener (staple) removal, and digital imaging (making use of Google's Open Source Tesseract OCR. Canopy's auto-classification automates 80% of records classification and with more training by company records managers, classification accuracy can reach upwards of 90%. Once classified, maintaining records is simple, with easy to use record retention and disposition policy management.

Ripcord Canopy is a SaaS solution based on AWS S3 storage that enables customers to quickly search and locate records with keyword, Boolean and filtered search. (Azure and Google options are also available).

Canopy allows customers to manage the full lifecycle of their enterprise content, whether it starts its life as a physical or digital document. It uses AI and ML to auto-classify and tag information within the PDFs.

Applications

Ripcord Workcells can process an entire 'banker box' within 1 – 2 hours whereas it would take a human about 4 – 6 hours to digitise the same content. At scale, Ripcord can process thousands of boxes per day. Using robots to digitise paper is inherently more secure. Robots do 90% of the paper handling and 90% of the indexing. This means there are many fewer people in the workflow that could access and read sensitive information.

The more places documents are stored (offices, file rooms and third-party facilities) the more challenging it becomes to enforce security and compliance policies. Ripcord Canopy integrates with existing identity management tools like Active Directory and LDAP, data is encrypted both in-transit and at-rest, and you can set access and retention policies for compliance.

Observations

Ripcord are often asked, "Why has no one done this before?" Creating a solution like this, one that converts the inherent chaos of paper inside boxes to ordered, digital documents that are easy to search, requires bringing together a variety of technical disciplines.

Ripcord has built a team, comprised of industrial roboticists, machine vision engineers, mechanical engineers, operations specialists, data scientists, machine learning experts, software developers, and infrastructure and security experts. The result is the world's most advanced paper handler. They continue to improve their technology and are currently working on a new solution that will be able to scan and read the contents of a box of paper records without opening it, available within 2 to 5 years.

During the introduction to Ripcord, we learned that the Dalia Lama reads TechCrunch and, after reading an article about Ripcord's technology, had recently initiated a conversation between the company and the Tibetan authorities to see if it might be possible to digitise around 30,000 sacred texts on very delicate rice-paper which is disintegrating and in danger of being lost to the world. Ripcord's high levels of corporate social responsibility has prompted them to respond very positively.

Companies presented

- CachePhysics
- Element Analytics

CachePhysics

(NB. Currently Rebranding)

Founded	2018 Headquarters: Redwood City, California
Founder(s)	Irfan Ahmad
Investment Funding	Currently in Stealth
Website	cachephysics.com
Sector	Autonomous performance management, Infrastructure
Key Points	• Cache monitoring and modelling software • Next-gen self-learning, intelligent caches • Empowering dynamic applications with 'workload-aware' performance SLAs and QoS guarantees

Overview

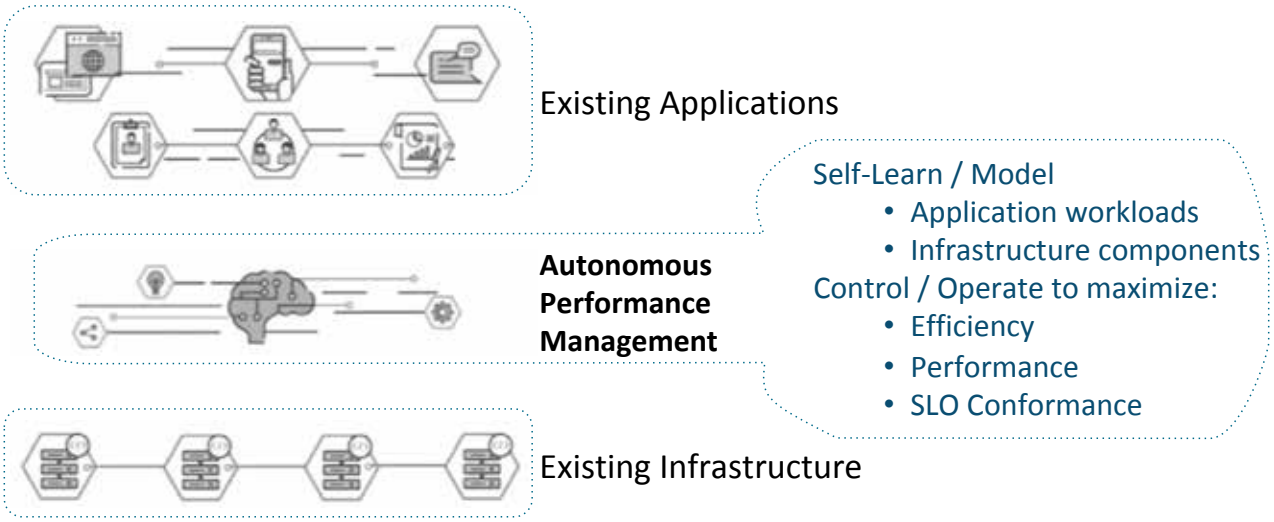
Irfan Ahmad, co-founder of CloudPhysics (who presented to the 2015 Tour at Polycom) has started a new venture called CachePhysics. He started his presentation this year by saying he needs to find another name and rebrand his 'stealth-mode' start-up. CloudPhysics provides predictive analytics for virtual datacentre management and have raised around $27.5M up to round C. Irfan has stepped-down from his role as CTO to concentrate on his new business which looks deeper into the datacentre infrastructure stack to assure today's dynamic applications with "workload-aware" performance SLAs and QoS guarantees.

Prior to CloudPhysics, Irfan was at VMware, where he was R&D tech lead and co-inventor for flagship products including Storage DRS and Storage I/O Control. Irfan worked extensively on interdisciplinary endeavours in memory, storage, CPU, and distributed resource management, and developed a special interest in research at the intersection of systems.

Technology

CachePhysics has set out to develop the world's first and only guaranteed latency quality of service (QoS) Autonomous Performance Management solution. They say that a 100ms latency increase impacts Amazon sales by a 1% drop. A 500ms performance hit causes Google to miss 20% clicks, massively impacting advertising revenue. Application QoS is difficult and costly. Today's continuous integration and continuous delivery cycles mean that weekly, daily and even hourly releases make traditional manual performance management infeasible. Access patterns can change in minutes and over-provisioning for peak demand is very costly. CachePhysics answer is to predictively allocate exact resources, remediate thrashing and hence eliminate unnecessary overspending. Their goal is that in less than a decade all cloud native applications will manage their own performance at the microsecond scale to provide autonomous performance management.

To achieve this, CachePhysics are developing an additional architectural layer that sits between existing infrastructure and applications that self-learns and creates a model of application workloads and the infrastructure components they consume. The model can then be used to control and operate the applications to achieve maximum efficiency and performance, ensuring service level conformance.



Their innovative software leverages self-learning and predictive algorithms to dramatically improve application performance:

- 5x improvement in data performance-efficiency
- 50% smaller cache infrastructure (DRAM) spend
- 500% reduction in cache performance incidents

Applications

CachePhysics provides auto scaling software for applications. CachePhysics uses patented self-learning technology to accelerate workloads (in database, storage system or application tier) by predictively 'right-sizing' and dynamically enforcing performance service level agreements as workloads and applications change dynamically.

They let developers write code to achieve business goals while trusting that the infrastructure will adapt to meet service levels for performance and cost. The cache infrastructure will self-tune and right-size throughout the application life-cycle thus ensuring the best performance at lowest cost. To achieve latency targets through modelling, CachePhysics predictively sets the operating points to hit latency targets at the lowest cost. As workloads shift or as applications evolve the solution can automatically adjust to hit service level targets at the lowest possible cost. "Apps can self-drive their performance".

Observations

High latency equals revenue loss. Autonomous performance management is a must-have for modern complex application architectures such as microservices. During his presentation, Irfan provided in-depth descriptions of the predictive modelling techniques that CachePhysics uses. The approach will need a lot of monitoring time and data to improve the model to an acceptable that ensures prediction accuracy. The team is engaged in private beta trials with several Fortune 500 companies, but they are keen to partner with other companies and requested introductions to delegates performance and DevOps teams.

Irfan's questions:

- Are there application performance management and cost reduction initiatives in your company?
- Are there plans to develop new or scale up existing latency-sensitive applications?
- Are there developers, DevOps or infrastructure engineers managing application caches (e.g. Redis, Memcached, etc)?

If the answer to ANY of these is YES, we would love to talk to you for 30 mins. This is NOT a sales call, we only want feedback. We promise it'll be a fun call.

Element Analytics

Founded	2015 Headquarters: San Francisco, California
Founder(s)	David Mount, Sameer Kalwani
Investment Funding	Total Funding: $22M Last Funding: Series A
Website	www.elementanalytics.com
Sector	Analytics, Big Data, Industrial, Industrial Automation, Internet of Things, Software
Key Points	• Creates industrial analytics software that empowers organisations to achieve new levels of operational performance

Overview

Element Analytics organises industrial data into an accessible and useful form and transforms it into actionable insights that help people make the best-informed decisions for greater worker safety, sustainability and business profit. Element connects data across the industrial enterprise to make it usable by people and applications, uncovering new insights.

Digital transformation provides a clear path for industrial organisations to increase profits and reduce operating risk but is exacerbated by the poor state of operations data availability and quality. The Element Platform is designed to do the heavy lifting of analytics data management and integration so that operations data is always ready and always trusted.

Industrial analytics requires aggregating, integrating and providing easy access to all the data generated from multiple industrial systems and applications, especially from time series data sources like process historians.

Technology

Aggregating data requires bringing together multiple data sources, applications, APIs and processes. Integration means joining many data types to provide the context critical to making the data analytically useful. Easy access makes data easy to explore for any user, feeding analytics and applications to uncover insights.

This requires a data lake to store the data and the ability to provide data context that defines the relationships between assets and the role each asset plays in day-to-day operations. To get operational data into this state and maintain it is prohibitively expensive and time-intensive to do manually.

Element is the only software solution that automates and continuously delivers a ready data lake for time-series and business data via Operational Digital Twins that can flexibly support any analytical activities.

Operating environments are constantly changing. Equipment is replaced, instrumentation is added or removed, process flows are adjusted. How can analytics tools accurately represent an environment that changes daily? The answer is Operational Digital Twins (dynamic digital representations of physical assets).

Maintaining and sustaining Operational Digital Twins in an ever-changing environment requires software to keep track of and manage these changes. The context around data (the relationships between assets and the role each asset plays in day-to-day operations) must be constantly streamed in and connected. This and more are all done in Element, which maintains a flexible Element Graph to store digital twins.

Applications

Element's out-of-the-box software solution empowers organisations to support their digital transformation strategy via a data lake and digital twins, reducing the time to analytics from months to seconds, and the cost by orders of magnitude. In Element, organisations can connect hundreds of operational systems/technologies, design and build digital twins, and keep them continuously up-to-date.

The platform has 3 key capabilities:

- Graph Studio (rapidly turn siloed data into an interwoven Element Graph)
- Data Integrity (ensure digital twins and data are continuously updated and accurate)
- Common Data Layer (turns data into an asset that all advanced analytics and BI tools can leverage)

Observations

Leading industrial organisations across verticals are choosing Element Analytics to enable their Digital Transformations.

Industrial companies are becoming increasingly sophisticated in their exploration of data analytics. They're aware of the promise of analytics but are increasingly focusing on getting down to business, encountering challenges and asking tough questions.

This means confronting cultural barriers ("this is how it's always been done") and working to ensure the broad availability of data and analytics (versus making them available only to an organisation's most motivated groups). They're also moving away from addressing narrow challenges with data and instead treating data as a powerful, broadly applicable and dynamic asset. To first-movers in the industrial sector, data is an asset on par with capital, equipment or personnel.

Greylock Partners

Founded	1965 Headquarters: Menlo Park, California
Founder(s)	Bill Elfers and Dan Gregory
Investment Funding	Investments in: Early Stage Venture, Late Stage Venture, Seed Number of Exits: 173
Website	www.greylock.com
Sector	Consumer, Enterprise Software, Venture Capital
Key Points	• Backs entrepreneurs who are building disruptive, transforming consumer and enterprise software companies • The original capital worth $10 million was committed by a group of six families

Overview

This was our second visit to Greylock Partners and once again they did not disappoint. Greylock are one of the longest established Venture Capitalists in Silicon Valley, having been established in Boston in October 1965 and relocated its headquarters to Silicon Valley in 2009. They invest from Seed rounds through Early Stage Venture to Later Stage Venture investments and now focus entirely on software. They view their 'sweet spot' as the first or second institutional investor Series A or Series B stages. Whilst their investments are split roughly 50/50 between Consumer Internet and Enterprise IT, they are still more comfortable with Enterprise IT, probably because of the backgrounds of their investing partners, many of whom are drawn from the Enterprise IT space.

The investing partners are active participants, they base this approach on the premise that "you invest in what you know" which enables them to understand what makes a good product. They want to be 'value add' investors, not passive, providing 'smart money'. Part of their evaluation criteria includes looking for teams that can scale and are capable of building durable long-term independent businesses.

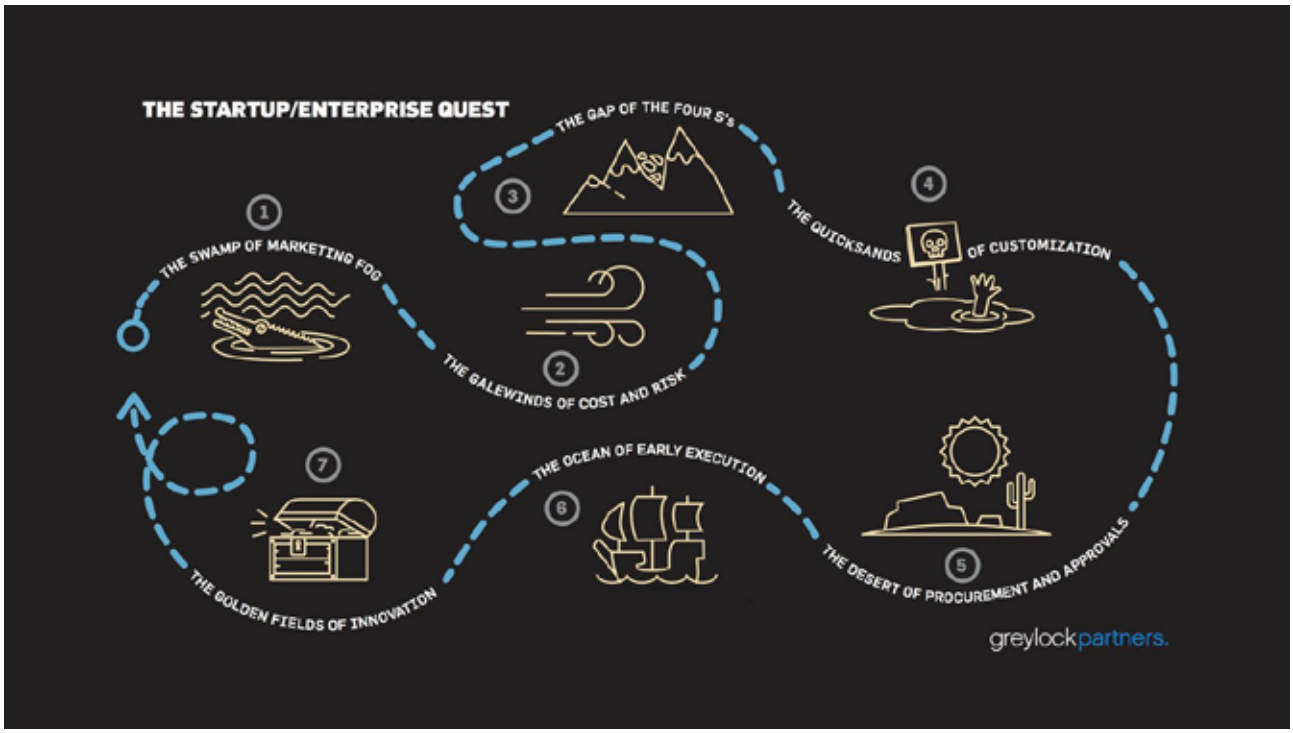
On the other hand, they also take the view that "contrarian and right is good, contrarian and wrong doesn't work" and emphasise that contrarian thinking is a good thing. Their investments are focused squarely in Silicon Valley, not international.

Their website states that Greylock Partners "backs entrepreneurs who are building disruptive, market-transforming consumer and enterprise software companies". They invest in companies that define new markets. Their portfolio includes some well-known names such as Airbnb, AppDynamics, Cloudera, Docker, Dropbox, Facebook, LinkedIn, Medium, Nextdoor, Palo Alto Networks and Workday. Some of them are listed on the New York Stock Exchange and have become 'Unicorns', companies with a market capitalisation of over $1 billion.

Their focus is on Consumer Internet and Enterprise IT. All Greylock partners are former operators and even today some continue to run large enterprises as their day job. They like to be known as the "nice VCs" and invest their efforts to ensure they maintain a culture of partnership with the entrepreneurs in whose companies they invest. They actively seek opportunities in areas where there is 'whitespace' or unsolved challenges.

One of the key success factors they (and other VCs like Sequoia) have identified in the past 2 to 3 years is the need for very good engineering and product design talent, a key topic being discussed and promoted in the valley now.

They have a strategy of being a long-term investor; their view is that a company needs to get to a certain scale of operation before starting to think about monetisation. They also see that some of their investments have disrupted and matured. For example, most advertising these days is undertaken through Facebook and Google, so they are starting to ask the question "is the age of social over?" However, Richards Research, a small consultancy based in New York, reported that millennials don't trust banks but do trust Facebook!



Companies presented

- Trifacta
- Delphix
- Rockset
- Rubrik

Trifacta

Founded	2012 Headquarters: San Francisco, California
Founder(s)	Jeffrey Heer, Joe Hellerstein, Sean Kandel
Investment Funding	Total Funding: $124.3M Last Funding: Series D
Website	trifacta.com
Sector	Analytics, Big Data, Information Technology, Software
Key Points	• Enables anyone to efficiently explore and prepare data by utilising machine learning • Trifacta leverages decades of innovative research in human-computer interaction, scalable data management and machine learning to make the process of preparing data faster and intuitive

Overview

Trifacta enables business users to take large, unstructured and complex data sets and transform them into structured, easy-to-analyse formats for analysis. Trifacta helps unlock the power of big data for use by the lines of business, which is very relevant for businesses that are already using big data systems like Hadoop.

We saw Trifacta just a year ago in 2017 but they have come a long way in a short time, so it was felt appropriate to let them re-present to this year's delegates.

Technology

Trifacta is a web-based system that optimises work by using 'machine learning' to assert data structure. Trifacta provides interactive, visual feedback so that workloads can be optimised on a very large data set with ease. The platform significantly shortens 'time to insight' by making data scientists and analysts up to 30x more productive, while also enabling dramatically better business decisions by giving business analysts direct access to Big Data.

Trifacta utilises a machine learning (ML) capability which it says, "gets smarter with experience and use". It sits between an enterprise's data storage and processing environments, and the visualisation, statistical or machine learning tools used downstream in the analysis process.

The process follow is a series of logical steps: discover, structure, clean, enrich, validate and publish.

Applications

Data transformation significantly enhances the value of an enterprise's Big Data by enabling users to easily transform raw, complex data into clean and structured formats for analysis. Leveraging decades of innovative work in human-computer interaction, scalable data management and machine learning, Trifacta says their technology creates a partnership between user and machine, with each side learning from the other and growing smarter with experience.

Observations

Businesses tend to transform data by passing it to 'data engineers' to restructure. Data is not self-service; rather it's a problem that's passed over to IT. Trifacta's mission is to move data maintenance into a self-service system and equip people with no technical backgrounds to interact with datasets in a productive way.

Trifacta works by wrangling and cleaning data. By cleansing the data to identify outliers, analysts can transform information-rich systems into self-service data sets for people who aren't engineers.

For example, if there's a retailer who wants to get precise forecasts for weather and how it will affect sales of certain products in stores across the country, they can begin by finding weather data, store location coordinates and opening hours and merge these different data sets using Trifacta. Originally, a person with a non-technical background would take this information and give it to IT but Trifacta cleans and wrangles the data into an easily accessible chart, eradicating the need for the 'IT middleman'.

Delphix

Founded	2008 Headquarters: Redwood City, California
Founder(s)	Jeffrey Heer, Joe Hellerstein, Sean Kandel
Investment Funding	Total Funding: $119.5M Last Funding: Series D
Website	www.delphix.com
Sector	Application Performance Management, Computer, Database, Virtualisation
Key Points	• Delphix's mission is to free companies from data friction and accelerate innovation • Dynamic Data Platform to connect, virtualise, secure and manage data in the cloud and on-premise environments

Overview

Delphix provides data virtualisation. They claim they help organisations release applications up to 10x faster by delivering secure, virtualised data across the application lifecycle. More than 30 percent of the Global 500 use Delphix software to deliver data across development, testing and reporting environments, improving developer productivity and data security on premises or in the cloud.

Delphix had previously presented to the Tour in 2017 but have moved on sufficiently in the last year and hence were given the opportunity to re-present to Tour delegates.

Delphix have seen four trends which they set out to address:

- The pace of change is accelerating
- Time to market, Governance and Efficiency as the key focus areas of many enterprises
- The move from on premise virtualisation to the cloud
- The move from waterfall to agile development

The four have been combined into a DevOps model that is being adopted by many organisations to enable continuous delivery. However, there is still a bottleneck in the data world and Delphix approach is to enable data to move securely at the 'speed of business'.

Technology

Trifacta is a web-based system that optimises work by using 'machine learning' to assert data structure. Delphix is a software-only solution that creates a fast, secure conduit to migrate, secure and provision data between data centres and clouds. Running as a virtual machine, Delphix synchronises with databases, files and file systems on premise and enables automatic, rapid provisioning of data to any point in time. Instead of migrating full volumes of data to hybrid cloud environments in the Elastic Compute Cloud (EC2), Delphix replicates only changed data, thereby enabling continuous updates while significantly reducing the costs of data transfer.

By providing self-service and APIs for integration with DevOps tools, Delphix further enables development teams to take full advantage of the benefits of elasticity and deployment flexibility in the cloud.

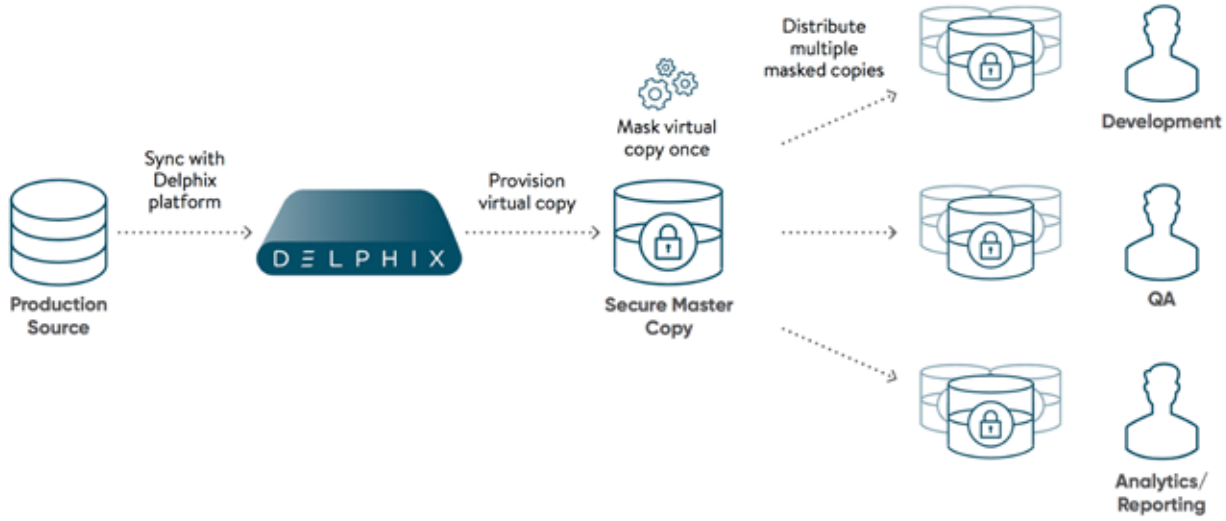
Delphix Data Masking is a software solution that gives businesses everything they need to identify, secure, audit and continuously protect confidential information. Delphix integrates the Delphix Masking Engine with the Delphix Virtualisation Engine to address the two key challenges that security-minded organisations face: masking data and then efficiently delivering it.

Delphix addresses the growing changes in software by providing a platform that allows companies to innovate more quickly and develop methodologies for writing code. With the rise of DevOps and the move from virtualisation into the cloud, there's an increased demand to innovate faster than ever before. Products are still developing at an inefficient pace, despite the fact that many companies now release new code into production several times a day. Data innovation is lagging behind the dramatic changes that have been made to infrastructure and there has yet to be a powerful change made to the ways we capture and deliver data.

Observations

Delphix Data Virtualisation decouples application data from physical hardware, allowing end users to access storage-efficient virtual data copies though a self-service model. This is achieved in three relatively straightforward steps:

1. Data virtualisation software is installed on-premise or in the cloud. The software is connected to production data sources which then collects data and stays synchronised with those sources as they change over time.
2. The data virtualisation platform serves as a single point of control for administrators to secure, archive, replicate and transform data. Administrators can instantly move data, automate processes and implement a self-service capability for testers, developers and other users whilst ensuring the data they are working on is secure and correctly governed.
3. The data virtualisation software allows users to provision fully functional virtual data copies that consume significantly less storage than the equivalent number of physical copies.



To assure and maintain data integrity and a full audit trail Delphix creates a full versioned timeline of the data by policy. The granularity of the audit timeline can go down to the minute and for Oracle databases can go down to the individual change record number for each change in the database.

Preserving data security and privacy is a key component of their offering, especially when releasing a virtual copy of a production database to a developer. Certain data may need to be masked, such as PII, HIPPA information, credit card numbers, social security numbers, etc. Their solution to this problem is Delphix Data Masking, which enables administrators to establish rules to mask data in a number of ways, including replacing an address in the same zip/post code with one that is not the data subject's real address.

The Delphix commercial model is to licence the files but, for example in the case of Oracle, users still need Oracle licences however they can give you much greater density than on a traditional Oracle landscape.

Delphix supports other databases, including Microsoft SQL, SAP ASE (Sybase), and DB2. Most recently they have OpenStack support plus KVM and Xen hypervisors as well as AWS and Azure.

Rockset, Inc.

Founded	2016 Headquarters: San Mateo, California
Founder(s)	Venkat Venkataramani, Dhruba Borthakur
Investment Funding	Early Stage, Seed Revenue: $2.4 Million
Website	rockset.com
Sector	Analytics, Cloud Data Services, SaaS
Key Points	• Formed by ex-Facebook data team. • Working on a fully managed fast data platform for enterprises • Provides instant live access to big data in the cloud

Overview

Rockset had previously been lined-up to present to us at Greylock in 2017 but were considered too early in their company formation to be able to provide sufficiently good account of themselves and their technology. They have since had a fruitful year working to deliver the "fastest path to apps and analytics with complex data". Their solution is due for launch in the summer 2018. We were given early access to and a demonstration of "how to put complex data to work" for data powered apps.

Technology

Data powered applications are the new normal. But real-world data comes from diverse sources and is often structured and semi-structured in nature. Data powered apps are difficult to build because current approaches to handling this type of data are broken. NoSQL databases need data denormalization and are difficult to query. On the other hand, 'sharded' SQL databases require an ETL pipeline, so they are not real-time in nature. Meanwhile, data warehouses are not built for operational speed queries and tend to have stale data.

Rockset is designed for building real-time operational applications on diverse, complex data sets using:

- Millisecond latency SQL
  Rockset maintains a live index that always stays in sync with multiple streaming (e.g.: Kafka) and static sources (e.g.: structured databases or AWS S3) for millisecond latency filters, aggregations and joins on multi-source data

- Schema-less
  Rockset auto-schematises semi-structured data, so nested JSON, Parquet, XML, CSV and TSV can be queried using SQL, without upfront data modelling or schema design. Rockset stores data in documents with all fields being fully mutable

- Strong Dynamic Typing
  In Rockset the data type is associated with the value of a field in each document to execute strongly typed queries on dynamically typed data, making data easier to work with in Python & Ruby, or C++ & Java applications

- Converged Data Indexing
  Rockset converges an inverted index and columnar index in a way that writes do not impact reads while optimising for key-value, document, search, aggregation and graph type SQL queries

- Serverless
  Rockset is completely serverless, so the minute you connect data sources to Rockset applications can start querying it.

Rockset is a Tailer-leaf-aggregator architecture (originally developed at Facebook for its social graph) and independently auto-scaling compute/storage make it one of the most modern and efficient data systems.

Applications

Rockset provides a real-time, schema-less, complex SQL (e.g. JOIN), converged Index, non-transactional database that supports hierarchical storage and burst processing in the cloud. With Rockset's distributed database system data can be left in-situ where it is with the indexes created in the cloud. This approach maintains the architecture principles adopted by their founding team who created the RocksDB, Hadoop.

Observations

Enterprises are still struggling to use complex data primarily because real world data is messy and cannot be used easily. Rockset are bridging the gap by changing the way data is stored, processed and accessed for making better, faster data driven decisions and data powered apps.

Rubrik

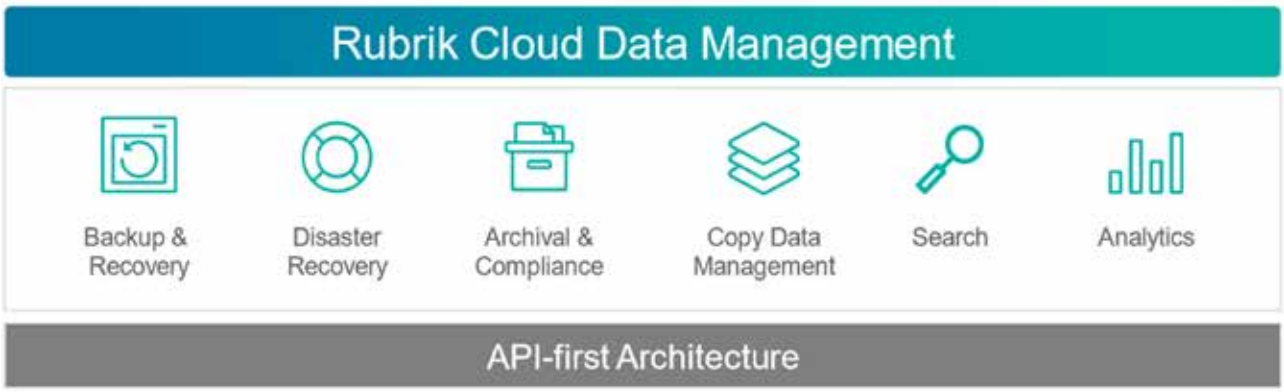
Founded	2014 Headquarters: Palo Alto, California
Founder(s)	Arvind Jain, Arvind Nithrakashyap, Bipul Sinha, Soham Mazumdar
Investment Funding	Total Funding: $292M Last Funding: Venture 1 Acquisition
Website	rubrik.com
Sector	Cloud Data Services, Cloud Management, Data Centre Automation, Security
Key Points	• Offers live data access for recovery and application development • Fuses enterprise data management with web-scale IT • Modern approach to backup and recovery

Overview

Rubrik has developed the world's first cloud Data Management platform for data protection, search, analytics, archival and copy data management for hybrid cloud enterprises. Enterprises use Rubrik to manage data at scale while realising data-driven services anytime, anywhere. Rubrik were previously presented to the Tour by Greylock in 2017.

Technology

A traditional backup approach adopted in many data centres was to copy data to an intermediate storage platform, copy it to tape, then ship the tapes through a service like Iron Mountain to a secure off-site store.



Rubrik offers an alternative, modern data management platform that unifies backup, instant recovery, replication, search, analytics, archival, compliance and copy data management in one secure fabric across the data centre and public cloud.

Applications

Rubrik provides backup and recovery, disaster recovery, archival compliance, copy data management and search and analytics all in a single abstraction layer through software. The product runs an 'API-first' architecture and allows users to search the cloud for stored data sets. The customer retains the cloud service account and relationship but gives Rubrik tenant access. All data is encrypted from Rubrik to the cloud and all data is sent with updates.

As well as utilising services on the public Internet, Rubrik also support Amazon's Snowball appliance. Snowball is a means by which Amazon ingest large data volumes by shipping storage appliances, the modern-day equivalent of "sneaker net".

Rubrik have gained FIPS142-2 level 2 compliance as well as compliance to ISO/IEC 15408 Common Criteria for Information Technology Security Evaluation, mainly referred to as common criteria. They will be Microsoft Azure compliant soon with Azure Stack, the Azure on premise solution, following shortly.

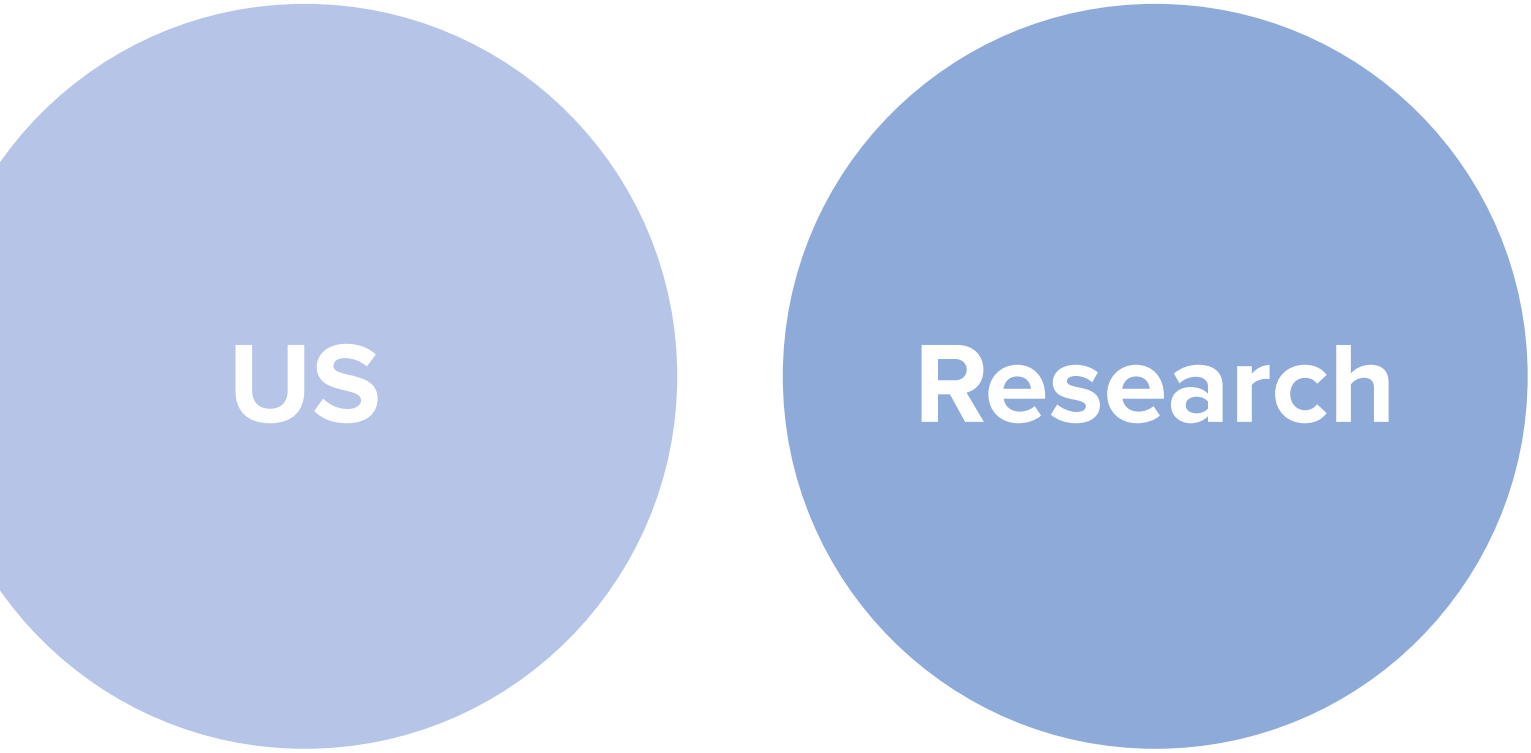
Observations

Rubrik is currently the fastest growing enterprise company in Silicon Valley and is overtaking every industrial vertical. Facebook, Honda, Tokyo Gas, and St. Jude Medical all use Rubrik to back up their data.

Rubrik was built with the cloud generation in mind: it shows you where data is being protected and provides automatic backups on any sort of virtual machine you can power it on. The software allows you to choose different levels of protection and the ability to map your Dropbox folder into your OS. Rubrik's multi-cloud functionality can build reports, compile summaries, and create galleries. Additionally, you can recover files directly and turn file management into a one-click process. With Rubrik, you can get an instant visual of the entirety of a file's history by simple opting to unfold it.

Rubrik provides a versatile policy in which its customers can start with data on premises and then eventually move over to the cloud. Rubrik is easy to get started; they claim it takes under two hours to run and deploy the entire system.

Wednesday 23 May



Tour

Ntegra Greenside

Knowledge for IT Leaders

NextWorld Capital

Founded	2009 Headquarters: San Francisco, California
Founder(s)	Craig Hanson
Investment Funding	Venture Capital: Early Stage Venture, Late Stage Venture 5 Exits
Website	www.nextworldcap.com
Sector	Analytics, Artificial Intelligence, Big Data, Cyber Security, Drones, Enterprise, Enterprise Applications, Enterprise Software, Finance, Financial Services, Internet of Things, Machine Learning, SaaS, Venture Capital
Key Points	• A venture capital firm that invests in enterprise technology start-ups • Lead Series A and B rounds and help entrepreneurs accelerate growth using their Global Enterprise Platform • NextWorld is headquartered in San Francisco with offices in London and Paris

Overview

Next World Capital is an international early-revenue stage venture capital firm investing in leading enterprise technology companies. The firm takes a focused, thematic approach with deep expertise in enterprise tech and specialised skills to help companies ramp-up during this rapid-growth period. Notably, NWC was the first U.S. VC to operate a comprehensive European Expansion Platform with strategy, sales and business development to help its companies access and grow in Europe. NWC is headquartered in San Francisco with additional offices in London and Paris.

They have built Next World's Global Enterprise Platform that has over 250 corporates using it. Their aim is to connect emerging technology companies with global opportunities. They achieve this objective in several different ways using a range of programmes they have developed to connect the two.

For young emerging companies NextWorld Capital have their Global Accelerator Program. This program helps start-ups with their strategy, go-to-market approach, finding and building partnerships, team building and most importantly customer introductions.

For Global Corporates they provide a programme to give access to the valley and start-up ecosystem, an innovation program and lastly, insight into technology and market trends in Silicon Valley.

Joining the two together are their thought leadership forums and their enterprise innovation centre, supplemented by a range of one-to-one consultations.

Their enterprise innovation centre is available at no cost with no contractual obligations and no commitment and can be used by multi or single company groups. There is no formal payment of membership.

Tom Rikert presented a selection of "Best of Breed" Silicon Valley start-ups (both in and out of the current Next World Capital portfolio) during the Ntegra day at the Stanford Park Hotel.

Companies presented

- Honeycomb
- ProsperWorks
- Mapper.AI
- Pilot.AI

Honeycomb.IO

Founded	2016 Headquarters: San Francisco, California
Founder(s)	Charity Majors, Christine Yen
Investment Funding	Total Funding: $15.5M Last Funding: Series A
Website	honeycomb.io
Sector	Analytics, Data Visualization, Database, SaaS
Key Points	• Honeycomb is a power tool built by engineers for engineers, designed to understand complex systems and answer high cardinality questions

Overview

Honeycomb provides full stack observability designed for high cardinality data and collaborative problem solving, enabling engineers to deeply understand and debug production software. Founded on the experience of debugging problems at the scale of millions of apps serving tens of millions of users, they empower engineers to instrument and query the behaviour of their system.

Anil, Honeycomb's head of strategy presented the view that "Nines no longer matter if users aren't happy". In a world where complexity is increasing, problems are becoming less predictable. With more interaction between people and systems and thousands of people "pushing the button", how can we make better engineering decisions, better priority decisions and better data insights?

Predictable	Unpredictable
Hardware failure	User behaviour
Physical/virtual exhaustion	Bot behaviour
Software resources	Interaction amongst multiple services
Service limits	Retry behaviour of multiple services
Known error states	3rd party API behaviour
Exceptions	Service degradation (not failure)
Counters and gauges	Cascading effects of faults / degradation

Honeycomb turns engineers into owners, enables them to ship faster and understand every event with 'Observability'. Playing on someone else's overused phrase, Honeycombs mantra is to "Move fast and fix things".

Technology

Complexity is increasing and problems are becoming more unpredictable. Nines don't matter:

- All the graphs are green, but the users don't care
- There are no alerts right now, but the users don't care

- No other users are reporting problems, but the users don't care
- 99th percentile looks ok, but the users don't care

...not this	...this
• Pre-canned answers to pre-canned questions	• Being able to ask any question
• High level trends and top 10 lists	• Being able to move from everything going on, to a single event and back
• Dashboards without context	• Rich content

• Grep	• Query
• Search	• Aggregate
• Graph	• Filter
• Data without context	• Data and context

• Proprietary agents	• Open source integrations and collectors
• Dictated data model	• Bring your own data model

Developers and integrators rarely know ahead of time what combination of attributes need to be correlated with some 'degenerate' behaviour. It could be a bad build ID or a combination of customer + replica set. With Honeycomb developers can store all the context without the cost of slow queries or impossible aggregations.

Honeycomb is a SaaS capability that accepts structured JSON objects with a write key (key value pairs). It is possible to submit events to the Honeycomb API, use their agent or integrations, instrument code with their SDKs or write your own custom integrations.

Applications

Honeycomb Observability solution is built for high-cardinality or sparse data. It has no rollups or aggregations. It provides access to raw events and traces, moving seamlessly between system wide events, individual traces and back. It has an easy to use query interface, has no indexes or schemas and provides high-performance querying and long-term event retention capabilities.

In control theory, 'observability' is a measure of how well internal states of a system can be inferred by knowledge of its external outputs. The observability and controllability of a system are mathematical duals. In the world of software products and services, observability means being able to answer any questions about what's happening on the inside of the system just by observing the outside of the system, without having to ship new code to answer new questions. When environments are as complex as they are today, simply monitoring for known problems doesn't address the growing number of new issues that arise. These new issues are "unknown unknowns," meaning that without an observable system you don't know what is causing the problem and you don't have a standard starting point/graph to find out.

Observations

Having an observable system means you have the instrumentation needed to understand what's happening in software. Observability focuses on the development of the application and the instrumentation needed, not to poll and monitor it for thresholds or defined health checks, but to ask any arbitrary question about how the software works.

It is not possible to predict what information needs to be known to answer a question that couldn't be predicted. Honeycomb is designed to gather as much context as possible from systems to assist with investigations and debugging new and complex problems in infrastructure.

Anil provided a good but very quick demo of the solution that provided insight into how Honeycomb works.

ProsperWorks

Founded	2011 Headquarters: San Francisco, California
Founder(s)	Chris Cheng, Jon Lee
Investment Funding	Total Funding: $87M Last Funding: Series C
Website	www.prosperworks.com | www.copper.com
Sector	Apps, CRM, Enterprise Software, SaaS, Software
Key Points	• Offers customer relationship management solutions for companies that use Google applications

ProsperWorks is now Copper (having rebanded after the tour).

Overview

Somewhere along the way CRMs became difficult to use. With a focus on design, usability and integration with Google's G Suite, ProsperWorks was built to deliver a seamless experience that puts email at the core of the workflow.

ProsperWorks eliminates manual data entry by automatically linking to all related emails, files and calendar events across the entire company. It keeps data up-to-date using PieSync's two-way capabilities so that any changes to Google contacts automatically update in ProsperWorks. It ensures no leads are missed by sending alerts when sales process isn't being followed to ensure leads aren't "falling through the cracks".

Technology

ProsperWorks is Google's recommended CRM. Many organisations use Google G Suite so they can have one system for everything related to their business. ProsperWorks is the CRM for G Suite, seamlessly integrating with Google Sheets, Docs, Slides, Calendar, Inbox and Gmail so using CRM is just as easy as the tools already in use every day. ProsperWorks CRM integrates all online tools in one place.

ProsperWorks was built using Google's material design principles so it looks and feels exactly like G Suite. This simple and clean interface makes navigating ProsperWorks easy and familiar. If you know how to use Gmail you'll know how to use ProsperWorks with no implementation or training required. ProsperWorks fits right into email so once the Gmail or Inbox Chrome Extension has been downloaded, users never have to leave their inbox.

Applications

ProsperWorks capabilities include:

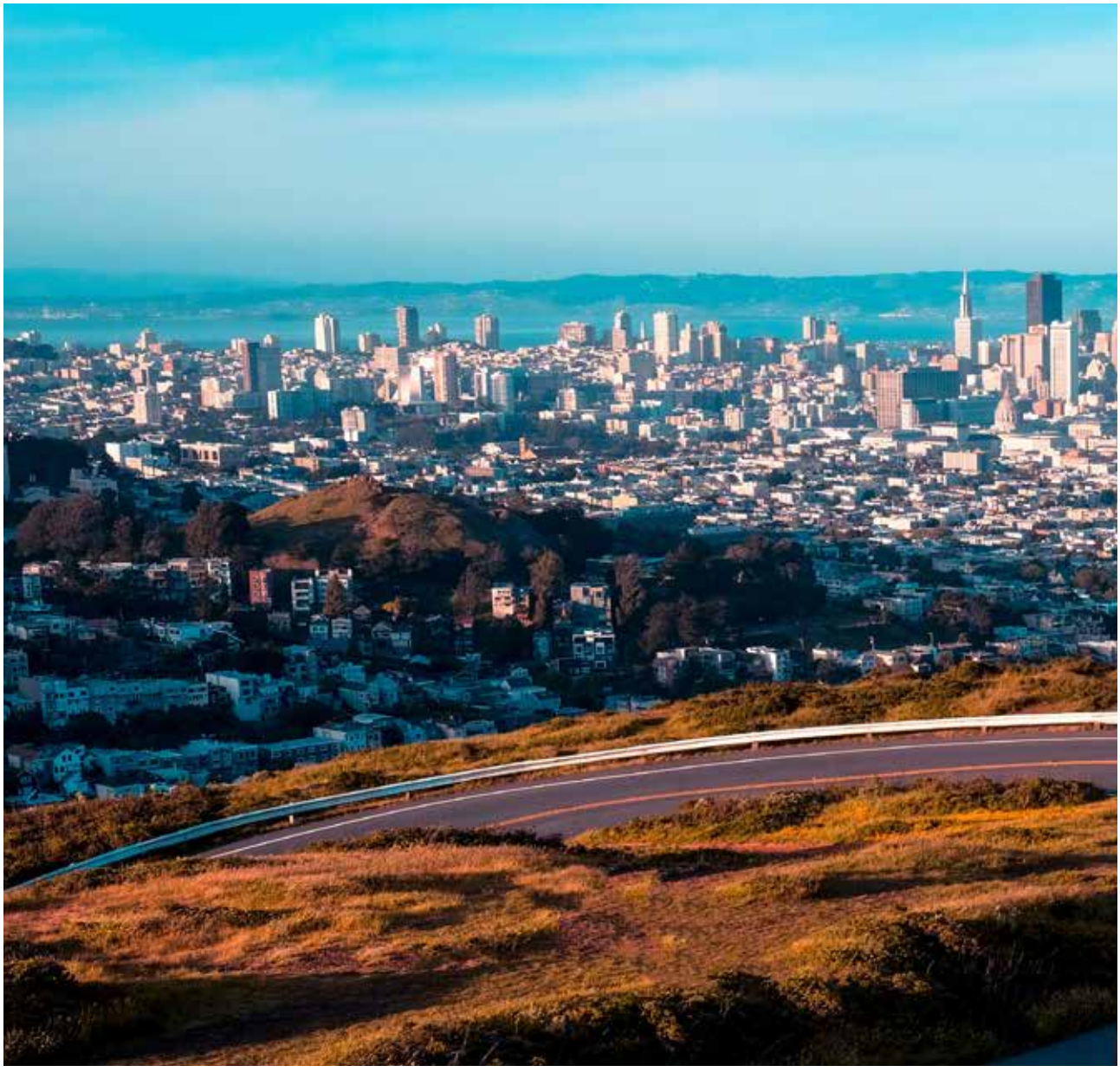
- Add a contact without touching the keyboard
- Track when emails are opened and viewed in real-time
- See recent activity from the inbox

It also has seamless integration with Hangouts Chat that ensures users never miss a beat on activity happening on contacts and accounts.

ProsperWorks integration puts a smart-bot inside the chat rooms users share with their teams. Think of it as a personal assistant for important CRM news, notifying critical changes made to Leads and Opportunities to ensure everything is moving forward.

Observations

CRM is a treasure-trove of crucial data for business. ProsperWorks make it easy to turn that data into insights with a Custom Report Builder that connects directly to Google Sheets. Export any dataset from ProsperWorks directly to Google Sheets and create advanced dashboards, reports, charts and graphs collaboratively within teams.



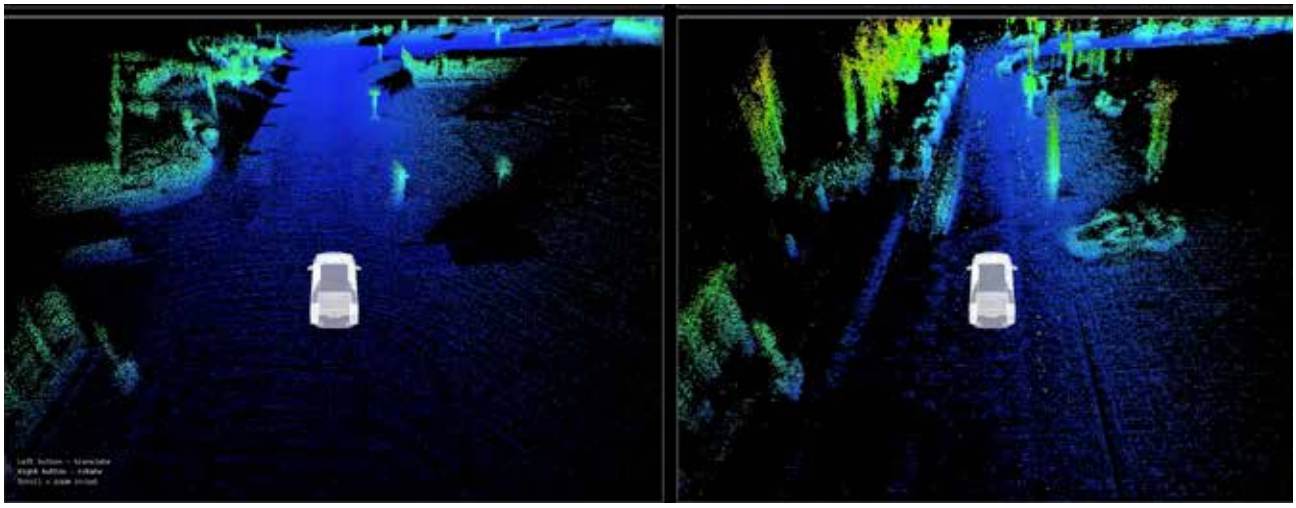
Mapper.AI

Founded	2015 Headquarters: San Francisco, California
Founder(s)	Alonso Patron, Nikhil Naikal
Investment Funding	Total Funding: $2.5M Last Funding: Venture
Website	mapper.ai
Sector	Software, Mapping, Image recognition, Geo Location
Key Points	• Provides Machine-Readable Maps (MRM) at global scale • Maps are created with the help of everyday drivers • Uses an in-house developed device and phone app • People across the globe can use their own car to begin mapping

Overview

Mapper is poised to become the largest data provider of Machine-Readable Maps (MRM) operating at global scale. They are actively growing four key teams rapidly: Research is responsible for algorithm development with a focus on computer vision; Engineering is designing and building a fully scalable cloud infrastructure and paired apps/tools based on the algorithms developed; Hardware is designing and building devices that collect and deliver the data required to meet their demands; and Logistics is focused on per-city rollout efforts.

LiDAR base map **Camera reconstruction**



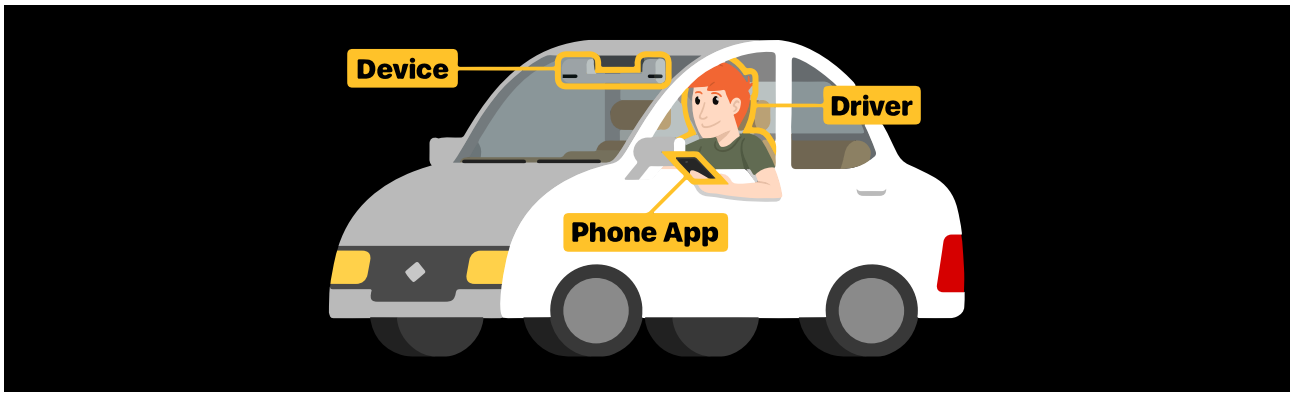
Technology

Mapper presents an on-demand solution that scales globally at a fraction of the cost of traditional mapping systems. Traditional mapping solutions use bespoke sensor arrangements on specially modified vehicle platforms, often using ultra-short-baseline differential GPS for high accuracy geolocation and geometry. The sensor array, often based on LiDAR needs regular re-calibration and maintenance, making the whole setup

prohibitively expensive to scale and deploy on-demand. Fully managed professional fleets are tied to fixed and contiguous geographies, providing low definition and infrequent updates.

Mapper enables customers to choose target location anywhere in the world thus providing global scale. Mapper simply sources freelance drivers on-demand and provides them with a device that is installed behind the driver's rear-view mirror. The driver is actively dispatched to drive the route six times (in both directions) and high resolution, high precision maps are built within 24 hours of collection and delivered to the customer.

The $300 camera-based sensor devices can generate LiDAR-quality maps with 10cm accuracy using image stacking techniques and interferometry. Image data is streamed to the cloud via an app on the driver's phone. Processing is undertaken in the cloud.



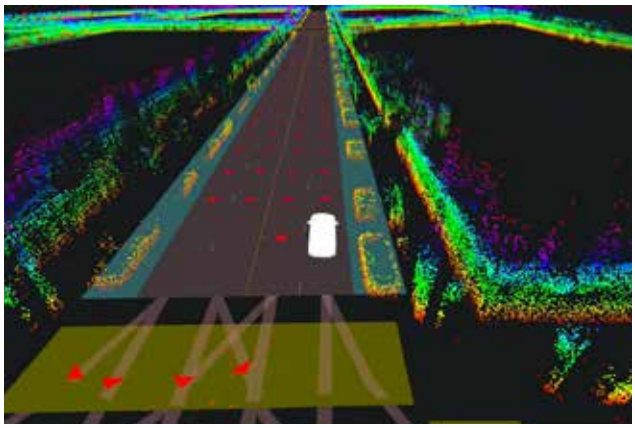
Applications

Mapper is the only accurate machine maps solution for all L3 & L4 autonomous vehicles which need machine maps for safe operation (L3 = driver assist/highway autonomy. L4 = fully autonomous). By 2023, two million vehicles will have L3 capabilities, as well as 50K L4 'robotaxis'. These maps are worth $400M (L3) and $50M (L4) annually and by 2030, machine maps will be worth $10B annually.

Observations

Thousands of people may soon be installing Mapper devices in their cars, hoping to make some extra money and in the process contribute to the next great crowdsourced project: a detailed and constantly updated map of the world's roads to be used by the swarm of self-driving cars that will populate them in future. The use of cheap commodity hardware, the economies of scale that can be achieved through mass-production, as well as the relentless march of Moore's Law, will see the price and accuracy of the Mapper devices fall significantly, making them very attractive to the 'Gig Economy".

A sample playback of some demo data is available here: http://bit.ly/2EeNuGs



Pilot AI

Founded	2015 Headquarters: Palo Alto, California
Founder(s)	Ankit Kumar, Brian Pierce, Elliot English, Jonathan Su
Investment Funding	Last Funding: Series A
Website	www.pilot.ai
Sector	Artificial Intelligence, Computer Vision, Internet of Things, Machine Learning
Key Points	• Software-based deep learning computer vision on commodity embedded platforms • Building a deep-learning based computer vision platform to solve real problems directly on compute-constrained embedded devices

Overview

Pilot AI are building a deep-learning based computer vision platform to solve real problems directly on compute-constrained embedded devices. They solve high-impact problems by developing computationally efficient algorithms to enable pervasive artificial intelligence. With their technology every device, no matter how big or small, will have intelligence. In a more intelligent world people and businesses are free from manual tasks and can make better decisions much more quickly.

Technology

Pilot AI solves high-impact problems by developing computationally efficient algorithms to enable pervasive artificial intelligence. With their technology every device, no matter how big or small, can have intelligence. In a more intelligent world people and businesses are freed from manual tasks and can make better decisions more quickly.

Much of the recent innovation within machine learning has revolved around scaling up models. These large models allow for more complex ideas and tasks to be captured at the cost of massively increased hardware requirements. Pilot AI has developed a suite of innovative algorithms that compress and optimise state of the art models to work with hardware that is more than an order of magnitude smaller than that used by competitors.

Pilot AI built everything from scratch from their databases to their deep learning framework. This means they have full control over every aspect of the Platform and can deliver customer-centric solutions unlike other offerings that utilise “one-size-fits-all” frameworks and methods. Using their proprietary artificial intelligence technology the Platform can run on commodity hardware to accurately and intelligently provide rich insights and contextual information ‘at the edge’, something that hasn’t been possible until now (other perception solutions just utilise outdated rules-based analytics).

Pilot models are trained on data that is relevant to customer needs and pain points, not just on the most common training datasets. Pilot’s team is always adding training data for new and existing applications, which helps models continuously improve. There are no hardware or firmware changes necessary to use the Platform. Because there is also no need to use the cloud, secure information stays secure and there are no additional costs associated with data transmission and processing.

Applications

Pilot AI provides perception solutions across a broad range of industries, from Retail and Smart Home to Government. The Pilot Platform is designed to be useful from the very start, from application-specific, out-of-the-box models to easy integration with edge devices and the cloud. The result is secure data, decreased data processing costs, and faster, more accurate insights.

Pilot AI is a drop-in neural network solution for computer vision applications. Their platform handles data ingestion, data labelling, model training and on-camera applications ranging from object detection to analytics packages that all run in real-time on even the most compute-constrained cameras.

The Pilot Platform supports most video formats, such as RTSP streams from standard IP cameras. The Platform utilises highly accurate, state-of-the-art computer vision models trained for specific applications. These models can be trained to detect general objects, like a person, or very specific objects (e.g. a brown cat). Using the Detection results the Visual Analytics modules generate further insights about the detected objects, such as the age and gender of a detected person. The Platform outputs information obtained from the Detection and Visual Analytics modules in a format that enables simple integration with any existing systems and tools. The Platform can run on any device with access to video data, which includes standard IP cameras and smart doorbells.

Observations

With the large-scale connection of terminal devices to the Internet, the massive amount of data being produced not only provides commercial value, but also presents a challenge to data processing. The need for real-time responses despite limited networking bandwidth, means that 'edge computing' has become a new development trend. There are many advantages of performing deep learning on integrated devices. They can perform direct computation without network connectivity and performing computations directly on integrated devices avoids the problem of latency. Storing data on terminals also solves many of the privacy concerns associated with cloud.

During his presentation Pilot AI co-founder and CEO Jonathan Su (a PhD in Computer Science from Stanford University who has expertise in data optimisation and high-performance computing) demonstrated a drone "Follow-me" application that uses a camera to track a person without the need for GPS tracking.

Other applications of Pilot AI that Jonathan demonstrated included the analysis of 'store traffic' in the retail sector and the detection of fatigue when using assisted driving systems that require the driver to stay alert and take-over control if an incident outside the car's operational capabilities occur.

In the long term computer vision will replace human "sight" in an increasing number of situations, freeing up large amounts of productivity, time and effort. In the future, it will be used in an increasing number of fields, making computer vision an interesting and lucrative market.

Thursday 24 May

US

Research

Tour

Ntegra Greenside

Knowledge for IT Leaders

NASA Ames Research Centre

Founded	1939 Headquarters: Washington, District of Columbia, USA
Founder(s)	Dwight D. Eisenhower
Investment Funding	NASA receives its funding from the annual federal budget passed by the United States Congress.
Website	www.nasa.gov/ames
Sector	Aerospace, Government, Information Technology, Innovation Management, Space Travel
Key Points	• one of 10 NASA field Centres • located at Moffett Field near Mountain View in the heart of Silicon Valley

Overview

NASA is responsible for the US civilian space program as well as aeronautics and aerospace research. For over 60 years Ames has led NASA in conducting world-class research and development. With 2500 employees and an annual budget of $900 million Ames provides NASA with improvements in: entry, descent and landing technologies; information technology; next-generation aviation improvements; astrobiology; airborne sciences; and small satellite programs.

NASA Research Park (NRP) at NASA Ames Research Centre is a world-class shared-use R&D and education campus for industry, academia, non-profits and government. NRP provides a physical place for innovation and entrepreneurship and serves as a technology accelerator through fostering both informal and formal collaborations. With commercially standard leases NRP houses over 70 industry and university partners collaborating with NASA on a variety of technologies and disciplines.

Hangar One at Moffett Field, built in the 1930s as a hangar for US naval airships, is one of the world's largest freestanding structures covering eight acres. The hangar is so big that weather such as fog can form inside just under the ceiling. Google has leased Hanger One and two other hangars capable of housing robotics research and space-exploration vehicles. The 60-year $1.16bn lease includes a $200m commitment to refurbish the hangars and improve the site with a museum and educational facilities. Moffet Field is situated next to Google's "Googleplex" headquarters and is used by Google to store its private jets, while a section of the land forms part of Google's expanding campus.

The NASA Ames Development Plan build out includes 4.2 million square feet of office, research and development space, university classrooms, labs, museums, housing and a conference and education centre.

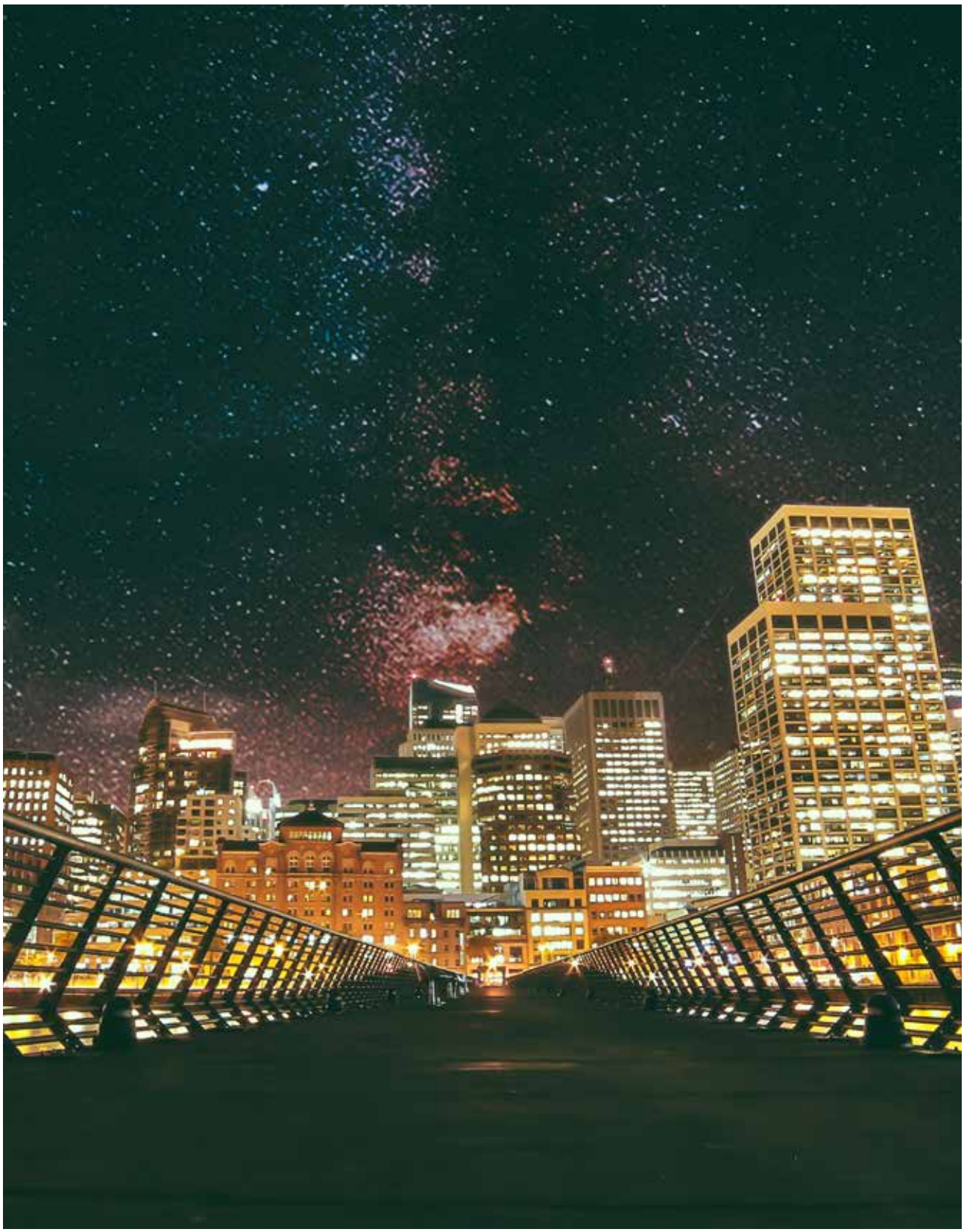
Observations

Our Thursday morning visit started with a presentation by Chief Scientist, Jacob Cohen, Ph.D., describing the formation of a "disruptive culture" and "thinking outside the box". This was followed by an introduction to the Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) programmes. These provide an opportunity for small, high-technology companies and research institutions to participate in government-sponsored research and development in key technology areas (contributing to NASAs missions and providing social benefit and economic growth).

Partnerships are a significant part of the NASA Strategic Plan. NASA recognises that they need to look for

innovative ways to partner to continue the development of new technologies for use in air, in space and on Earth. To fulfil this strategy, Ames Research Centre Partnership office funds research, development and demonstration of innovative technologies, develops technical partnerships with non-traditional space nations and works with commercial space companies to identify NASA technology that may be useful to them and to infuse commercial technology into NASA missions. It transfers NASA technology to the private sector through licenses and agreements.

See “Photo Ops” page of the NRP Post:
https://www.nasa.gov/sites/default/files/atoms/files/nrp_post_summer_2018.pdf



Highland Capital Partners

HIGHLAND
CAPITAL PARTNERS

Founded	1988 Headquarters: Palo Alto, California
Founder(s)	Bob Higgins, Paul Maeder
Investment Funding	Early Stage Venture, Late Stage Venture, Seed Number of Exits: 90
Website	www.hcp.com
Sector	Finance, Venture Capital
Key Points	• Global venture capital firm focused on "putting the entrepreneur first" • Offices in Silicon Valley, Boston and Shanghai • Highland has raised over $3 billion in committed capital and invested in more than 225 companies

Overview

Highland Capital Partners are a global venture capital firm focused on putting the entrepreneur first. With offices in Silicon Valley, Boston and Shanghai, Highland has raised over $3 billion in committed capital and invested in more than 225 companies resulting in category-defining businesses across consumer and enterprise technology. Investments include 2U, Ask Jeeves, Bromium, Gigamon, Leap Motion, LevelUp, Lycos, MapQuest, Nebula, QD Vision, Qihoo 360, Quattro Wireless, RentJuice, Rent the Runway, Starent Networks, Sybase, Violin Memory, VistaPrint and WePay. Highland also serves as a mentor for future entrepreneurs through Summer@Highland, a program that provides students and university-affiliated start-ups with the environment and resources to take their initiative to the next level.

vArmour

Founded	2011 Headquarters: Mountain View, California
Founder(s)	Michael Shieh, Roger Lian
Investment Funding	Total Funding: $83M Last Funding: Series D
Website	www.varmour.com
Sector	Mobile, Network Security, Security, Software
Key Points	• Offers data defined perimeter security solutions for mobile, virtual and cloud platforms • Founded by security and technology experts from Juniper, Netscreen, RSA and Silvertail

Overview

The vArmour cybersecurity showcase was hosted by Highland Capital Partners. Keith Stewart from vArmour provided an introductory briefing on all things cyber' sstating that "as apps are becoming more and more complicated and stretched, vulnerabilities and exploits are becoming more and more complex to identify and mitigate. This 'perfect storm' is being exasperated by the difficulties young and innovative cyber companies have breaking into the marketplace. vArmour like to support new cyber companies with financial support and mentoring through the tricky areas of compliance, risk, DevOps and cloud, in what's become a 'zero trust environment'. vArmour is a data centre and cloud security company that delivers software-based segmentation and micro-segmentation to protect critical applications and workloads through 'the industry's first' distributed security system. Based in Mountain View, CA, the company was founded in 2011 and is backed by top investors including Highland Capital Partners and Menlo Ventures.

The vArmour Distributed Security System (DSS) is deployed across the world's largest banks, telecom service providers, government agencies, healthcare providers and retailers. Partnering with companies including AWS, Cisco, HPE, and VMware, vArmour builds security into modern infrastructures with a simple and scalable approach that drives agility and operational efficiency.

Technology

vArmour is a leader in data defined perimeter security. Where the world of mobile, virtualisation and cloud has moved beyond the traditional perimeter, threats are more sophisticated and CIOs require greater visibility and control of their data.

The recent worldwide occurrences of ransomware attacks are an example of how reliant enterprises are on data that is stored across multiple system silos. The lateral movement of a virus or malware infection results in contagious spread of infection very quickly.

Challenges and pain points with traditional cyber-security models, include:

- Fragmentation of data across on premise and cloud services
- Data security challenges introduced through cloud transformation

- Regulation and compliance (e.g. GDPR and NIST breach disclosure)
- The arms race bought about by the continuous cycle of security refresh

vArmour's Distributed Security System (DSS) protects what matters by securing and segmenting apps and users, not the network protocols and VLANs. Their solution enables the deployment of secure applications faster through simplification, policy management and operations automation. vArmour can be deployed on any infrastructure resulting in one fabric and one policy infrastructure to deliver compliance.

Observations

Hardening perimeters has become an outdated security measure, even though most companies spend 80 to 90 percent of their web security budgets on firewalls and perimeter security. It has become increasingly difficult for companies to prevent malware from infecting environments. Once security has been breached, malware moves laterally across systems in an autonomous 'contagion' which can compromise sensitive information.

vArmour helps people with either their existing data environment or new cloud environments by wrapping and protecting critical data. They describe their solution as the security for the new IT; a protection that is oriented on critical applications of data.

vArmour is usually introduced during difficult periods for a company, e.g. when an organisation embraces some form of cloud, or when they attempt to acquire better regulation and compliance processes, or try to re-evaluate costly firewall solutions.

vArmour suggests that systems should deploy a first and last line of defence in which a set of controls are enforced around the most important assets. vArmour describes their system as a "security fabric" that runs across all physical, virtual, and cloud assets. With vArmour, companies can stop caring where the asset is and start caring about what it does instead.

vArmour's system builds an additional set of technology by learning a network's existing environment and applying rules with which to abide the structure that's been put into place. Their recommended approach to companies is that they take on a form of "front door control." Most networks have multiple layers of systems and vArmour argues that not only are these layers painful and ugly, but they don't provide security. Instead, vArmour offers a control layer of security that establishes an integrity of control, continuously testing and validating a company's most critical and sensitive information.

Companies presented

- EnVeil
- Okera
- Templarbit
- Hyas
- Preempt
- Cycognitio

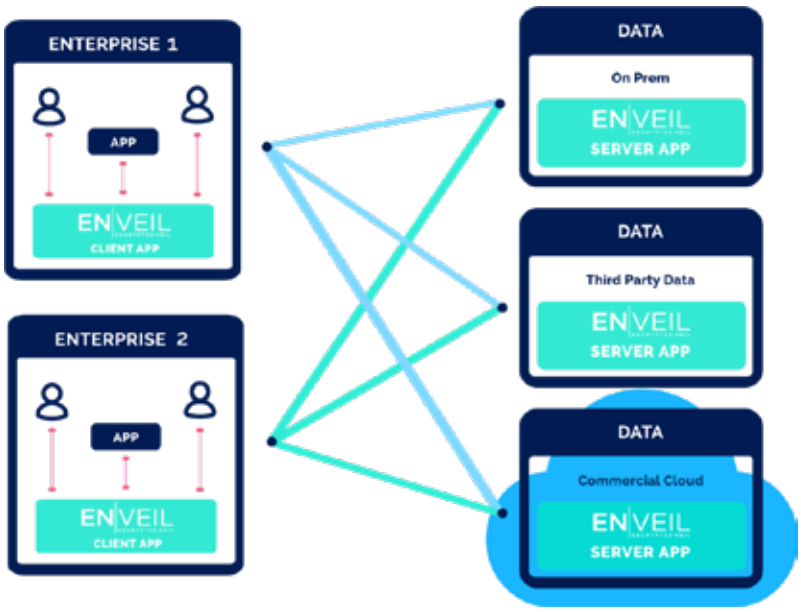
ENVEIL

Founded	2016 Headquarters: Fulton, Maryland, United States
Founder(s)	Ellison Anne Williams
Investment Funding	Total Funding: $4M Last Funding: Corporate Round
Website	www.enveil.com
Sector	Analytics, Cloud Data Services, Security
Key Points	• Uses homomorphic encryption to protect data interactions including search and analytic queries • Complimentary to encryption at rest and in transit • Provides trusted compute in untrusted areas

Overview

ENVEIL (Encrypted-Veil) has developed a homomorphic encryption based solution that enables encrypted data to be processed in its encrypted state, removing the need for data to be processed in an unencrypted state where it is most 'vulnerable' to being accessed or compromised. Enveil has solved the issue of protecting data in use using core technology that was developed inside the US Government's National Security Agency.

ENVEIL is the first commercial solution that provides a scalable framework to enable enterprises to securely operate on data without ever revealing the content of the operation, the results, or the data itself. Nothing is ever decrypted during the entire processing lifecycle. Analyse, search and perform calculations on sensitive data without ever decrypting anything, eliminating the risks of theft or incidental exposure. What was once only theoretical is now possible with Enveil.



Technology

Enveil's ZeroReveal solutions close the last gap in data security by protecting data while it's being used, the 'holy grail' of data encryption. They are the first and only commercial solution to ensure this full lifecycle security at scale achieving previously impossible levels of data security. Extracting value from data by performing actions such as search and analytics has historically required decryption, creating critical points of exposure. It's far too easy to assume current security practices already have this covered - They don't..

Enveil use homomorphic encryption techniques to solve this problem in a way no one else can. Whether performing searches or analytics on data, seeking information from a third-party data provider, or driving revenue by securely monetising data assets, Enveil ensures nothing is ever revealed during the entire processing lifecycle. The content of the interaction, the results and the data itself, are always protected so users can interact with data at scale in a completely private and secure manner.

Applications

- Trusted compute in untrusted locations
- Military-grade encrypted search to extend the boundary of the enterprise's trusted compute into untrusted locations
- Non-intrusive

  The capability does not require any changes to a system architecture, data storage format or technology, or application code

- Lightweight API-based architecture

  Enveil Client and Server applications function as a point-to-point proxy layer to perform secure computations

- Customer keeps their keys
- Keys never need to leave the owner's custody even when processing data outside their walls

Enveil reduces liability for companies processing datasets that include in-scope GDPR personally identifiable information (PII) data, or utilising third-party data services that include PII data. By implementing ZeroReveal Search companies can reach compliance without overhauling their current practices and resources.

Some datasets are more valuable than others and organisations must ensure that their most sensitive data is protected at all times. Enveil's ZeroReveal Compute Fabric eliminates risk of exposure by never decrypting an organisation's 'Crown Jewel' data during processing.

Enveil's Never Decrypt security posture allows organisations to process their most sensitive workloads and data in the Cloud, extending the boundary of the enterprise's trusted compute perimeter. Users can securely operate on both encrypted and unencrypted data in the Cloud, on prem, or anywhere in between.

Observations

Founded by U.S. Intelligence Community alumni with backgrounds in mathematics, algorithmic and machine learning, Enveil is revolutionizing data security by addressing a 'Data in Use' vulnerability that has challenged the us for over 20 years. The challenge of encryption is that at some point data must be decrypted in order to process it. Enveil has an answer to the encryption-in-use problem. Its ZeroReveal Compute allows searches or analytics to be performed on massive amounts of encrypted data without ever decrypting it.

Okera

Founded	2016 Headquarters: San Francisco Bay Area, West Coast (US)
Founder(s)	Amandeep Khurana, Nong Li
Investment Funding	Total Funding: $14.6M Last Funding: Series A
Website	www.okera.com
Sector	Cloud Data Services, Software
Key Points	• Enable the management of data access and governance at scale • Enterprise-wide platform facilitates provisioning, accessing, governing and auditing of data

Overview

Okera claim to be the first software provider to enable the management of data access and governance at scale for today's modern heterogeneous data environments. Okera's Active Data Access Platform allows agility and governance to co-exist and gives data producers, consumers and stewards the confidence to unlock the power of their data for innovation and growth. This unique, enterprise-wide platform facilitates the provisioning, accessing, governing and auditing of data in today's multi-cloud, multi-data-format and multi-tool world.

Data is the most valuable asset businesses have to drive innovation. Okera's platform tackles the hardest issues behind data access and governance across distributed environments, "giving you the ability to explore your data's potential like never before".

Okera opens data for greater innovation by scaling access and governance across heterogeneous, distributed data environments. The Okera Active Data Access Platform manages data access across a multi-cloud, multi-datastore and multi-tool world reducing friction between agility and governance. With greater accessibility, protection and visibility, customers have the confidence to move forward to innovate.

Technology

The Okera Active Data Access Platform is a unified platform for actively managing data access.

Data Access Service: Easy access. Fine-grained control.

Consumers only see data they have access to and data is abstracted from the original source and not copied:

- Handles I/O and provisions data to analytics tools after applying schema, fine-grained access policies and other transformations (UDFs, tokenization, masking etc.) with high performance
- Data provisioned in a format that is easily consumable
- Multiple tools like Spark, Python, SQL engines, Notebooks and business intelligence tools like Tableau and Excel are supported

Schema Registry: Faster Access, Greater Collaboration

Provide a unified view of all relevant data and user access across multiple analytical tools:

- Metadata representation for easily understanding, finding, and querying data
- Schemas, dataset sizes and who owns them, tags, annotations, basic quality metrics
- Access via the Hive Metastore API, REST APIs or a GUI

Policy Engine: Scalable Protection, Greater Agility

Allow data teams to define and manage data access policies at scale including:

- Fine-grained access control
- Auditability, tokenization, and anonymization of data on the fly
- Supported at several granularities: database, dataset, rows, columns and even cells

Audit Engine: Better Visibility

Track data access, and understand how your data is being used at a detailed level through a rich audit dashboard including:

- Data usage tracking by user
- Most common datasets
- Most commonly used tools

Applications

- For consumers
  It is easier to find and access data quickly using a variety of analytics tools

- For stewards
  It is easier to define, enforce and audit access policies at a fine-grained level

- For Owners
  It is easier to publish datasets and get visibility into how the data is being used

The EU's General Data Protection Regulation imposes strict guidelines when collecting, maintaining, and using consumer data. Okera provides the necessary controls to not only adhere to these guidelines but also set businesses up for future success. The more you know, the better. With Okera's catalog services data owners or stewards create access policies to datasets so that various teams can perform analytics or build new applications and data pipelines. Okera provides the ability to tag PII data so that businesses can pinpoint their specific GDPR exposure.

Observations

Innovation is happening faster than ever and it's essential that data access is given to the people who need it most. Yet understanding who is using it and how it is being used is always of critical concern.

With Okera, agility and governance doesn't have to be an "either / or" decision. With greater visibility and control you can unlock the potential that data has to drive innovation.

Templarbit

Founded	2017 Headquarters: San Francisco, California
Founder(s)	Bjoern Zinssmeister, Matthias Kadenbach
Investment Funding	Total Funding: $3M Seed
Website	www.templarbit.com
Sector	Apps, Cyber Security, Information Technology
Key Points	• protects applications from malicious activity known as “xss attacks” • intelligent security platform that thinks “like the bad guys”

Overview

Templarbit protects applications from malicious activity through use of Cross Site Script (XSS) attacks which are usually delivered in a browser side script and other common vulnerabilities. XSS attacks are one of the top ten methods that the Open Web Application Security Project (OWASP) identify as the most critical web application security risks that organisations should test for during code QA and protect against. Templarbit’s solution protects against all the OWASP Top 10.

Technology

Most breaches use the application layer as the attack vector. Templarbit stops these attacks with an application security platform that is intelligent, context aware and easy to deploy.

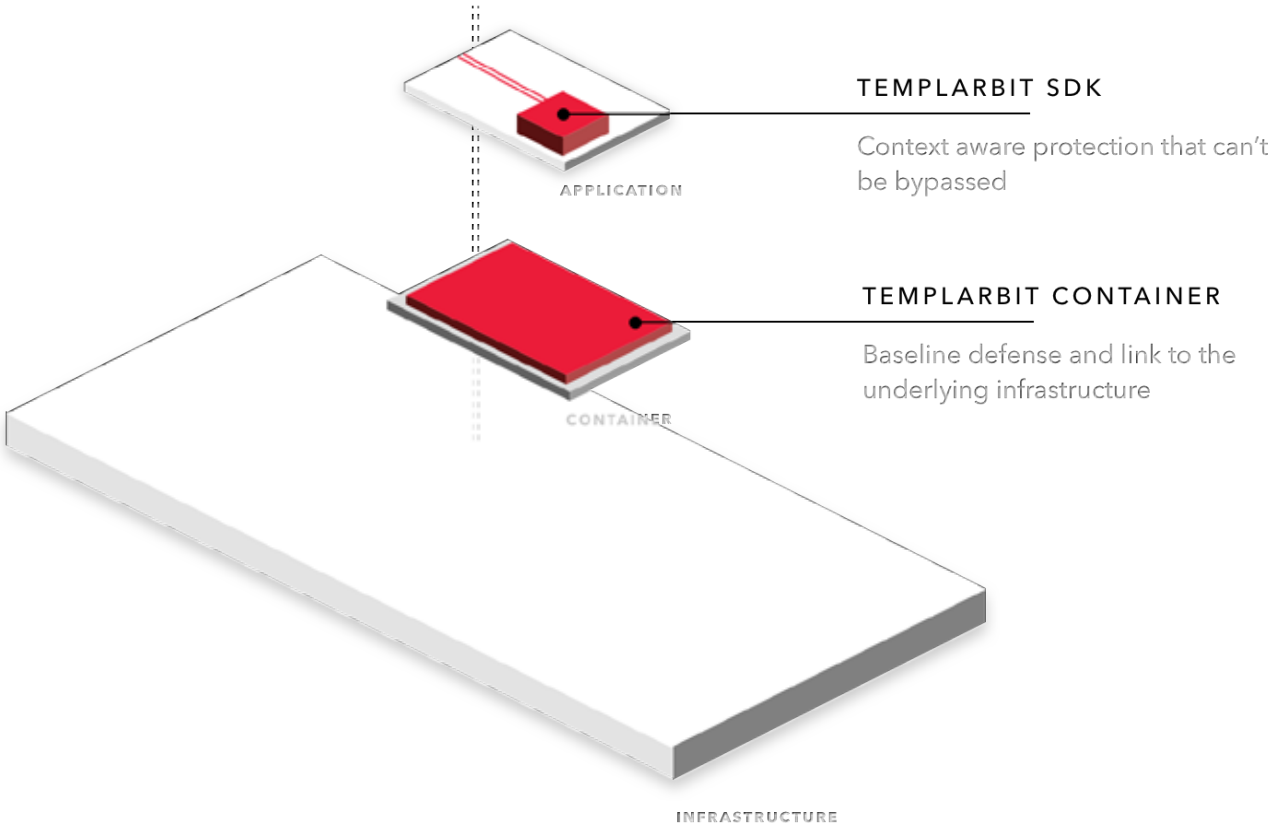
To monitor and defend the complexity of software applications continuously at scale is a challenging task for humans. Templarbit uses artificial intelligence to analyse data and build up scalable defences. Powering things behind the scenes are a combination of our proprietary data and machine learning models. This allows Templarbit to automatically make a decision on policy changes without relying on human input.

Templarbit can be easily deployed into an application stack:

- Add the SDK into the web application or API
- Deploy the Data Agent with the data store
- Run the Container inside the cluster

or deploy all three options to have complete visibility and control.

Once added to the stack, Templarbit will equip the system with defence capabilities that protect from breaches. The data passed through the Application Layer will be analysed on-premise and all signals of malicious behaviour will be reported back in real-time.



Applications

- Spot unusual behaviour that shows signs of a breach in progress
- Provide security teams with greater insights into areas that are likely vulnerable
- Analyse policy violations and automatically adjust whitelist items without human input

Observations

Templarbit is one of the most direct ways to improve a company's security posture and will help ensure that teams can stay on top of things by surfacing security threats in a format everyone can understand.

Hyas InfoSec

Founded	2015 Headquarters: Victoria, British Columbia, Canada
Founder(s)	Christopher Davis, Jeff Spencer, Sasha Angus, Steve Heyns
Investment Funding	Seed
Website	www.hyas.com
Sector	Cyber Security, Network Security, Security
Key Points	• Threat research, intelligence, compromise detection, and last leg attribution • Protection through better attribution and remote compromise detection at a global scale

Overview

Hyas are a threat research, intelligence, compromise detection and last leg attribution solution. Hyas's Comox solution is an AI enabled cyber security investigation suite that trawls numerous data sources including dynamic DNS records It then uses the massive data set and data science analytics to rapidly connect specific attack instances and campaigns to historical Indicators of Compromise (IoC). It enables security analysts to research both malware and their corresponding network traffic to pinpoint sources of malware and therefore enable a CSOC to respond to them. Hyas Comox feeds into an enterprise Security Information and Event Management (SIEM) solution.

Technology

Comox by Hyas uses exclusive data sources and non-traditional collection mechanisms to provide a powerful investigation and attribution system. Comox gives analysts, researchers and investigators the tools to be much more effective at their jobs, vastly increasing the accuracy of their findings. With an easy to use Web UI and JSON API, Comox ties together billions of data points into an incredibly powerful research and attribution platform tracking adversaries to their doorstep.

Comox lets users connect specific attack instances and campaigns to billions of historical indicators of compromise faster than ever before. It helps users research millions of malware samples and their corresponding network traffic and can help bring invaluable insights into existing SIEM efforts.

Applications

Comox acts as a 'force multiplier' for institutions which are having to work harder to protect themselves from breaches that can cost millions. For these companies, Hyas technology makes their security operation centres much more effective, vastly increasing the accuracy of their findings and allowing them to quickly respond to potential threats.

Comex can link an IP address to a location with great accuracy. When Hillary Clinton received an email as secretary of state, the message landed on a server registered to her home and physically located in the area, according to Internet registration records. The unusual arrangement for such a high-ranking official was first

reported by the Associated Press. The Wall Street Journal confirmed these findings by consulting with Hyas Technology, which researches who controls servers and Internet addresses used in computer intrusions.

Observations

Hyas believe in the internet as an intrinsic force for good and use that philosophy as the basis for their company. As a platform for change and innovation, the internet is unparalleled. Hyas stand firm on the right of individuals, societies and companies to embrace the internet's potential free from fear of harm, political retribution, or censorship. They say they offer uncompromising protection for their clients as an expression of their uncompromising support of free speech.

Preempt Security

Founded	2014 Headquarters: San Francisco, California
Founder(s)	Ajit Sancheti, Roman Blachman
Investment Funding	Total Funding: $10M Last Funding: Series A
Website	www.preempt.com
Sector	Cyber Security, Network Security, Security
Key Points	• Protects organisations by eliminating security threats • The Preempt platform pre-empts threats by continuously adapting according to identity, behaviour and risk • Ensures security threats and 'risky' employee activities are responded to with the right level of security at the right time • The platform scales to provide comprehensive identity-based protection across organisations of any size

Overview

Preempt protects organisations by eliminating security threats. Threats are not black or white and the Preempt Platform is the only solution that pre-empts threats with continuous threat prevention that automatically adapts based on identity, behaviour and risk. This ensures that both security threats and risky employee activities are responded to with the right level of security at the right time. The platform easily scales to provide comprehensive identity-based protection across organisations of any size.

Technology

Preempt Security combines elements of identity (who are you? an employee? executive? contractor? service account?) with behavioural analysis (what are you doing? is this normal relative to a baseline of your typical behaviour?) and risk scoring (how sensitive is this asset?) to help make more effective and more nuanced security policy decisions than the standard binary choice of allow or block. The goal is to make traditionally static identity and access management (IAM) more dynamic and better able to deal with threats such as compromised credentials.

Historically, IAM has been regarded as a separate discipline, often reserved as a tool for onboarding and offboarding employees, and effectively cordoned off from the 'real' security work done by network and endpoint security tools. With some exceptions (e.g., remote network access via VPNs or for access to sensitive servers or applications) user identity has been largely ignored as an input into the access decisions made by network security devices such as firewalls. Yet there are logical reasons why identity should be relied upon more directly for security decisions and policy implementation, particularly because the rise of cloud and mobile computing have undermined the traditional notion of a hardened perimeter between 'trusted' and 'untrusted' resources.

The Preempt Platform is an 'adaptive threat prevention platform' that consists of several components:

- Behavioural Firewall
- Preempt AnyApp
- Preempt Insights
- Threat Hunter
- Preempt Inspector

The idea is to combine elements of identity with behavioural analytics and risk scoring to help manage and control access to resources and, in the process, helping to prevent and respond to threats.

Architecturally, Preempt uses a proxy design, with a virtual machine image that sits on the network and receives and forwards traffic to and from domain controllers. Preempt is developing its technology to also sit in front of IDaaS platforms such as Okta, Ping and OneLogin, or VPN gateways as enforcement points for gating access.

The platform can be configured in active mode for real-time identity verification, or in out of band for near-real-time actions. The platform takes in information such as network traffic and activity logs from firewalls, VPNs, single sign-on (SSO) portals, cloud applications, SIEM tools and other data sources to create a baseline for detection of anomalous behaviour.

Applications

By operating at the network layer, Preempt can provide a view of all authentication and access events across the enterprise to enable a more flexible and granular set of access options beyond the standard 'allow completely or block fully.' The intent is to allow enterprises to address key use cases such as preventing compromised credentials, internal threats and privileged access abuse.

The Preempt Platform analyses a variety of factors including role, password attributes (including password strength) user behaviour and association with risky entities to generate a risk score for each user that can be used to make access decisions. Once a user score is generated, customers have a variety of enforcement options to respond to threats (Preempt's Behavioural Firewall). These include sending an email or SMS alert, forcing a password change, asking users to reauthenticate, stepping up to MFA, isolating the device or blocking the transaction completely. Preempt Insights provides analytics, visualisation and reporting to help identify risky users, devices or accounts. Threat Hunter is a querying tool that allows analysts to search on a variety of criteria to help identify unknown risks or for forensics purposes.

Observations

An advantage of Preempt's architecture is that it works for SaaS apps, custom apps running in a private cloud or IaaS, and on-premises apps that have historically been difficult to set up with adaptive MFA. Preempt Any App serves as an LDAP or Kerberos proxy that will intercept any application access attempts and integrate with any existing MFA product at the network layer, providing secure access without the need to make any changes to the application code.

Preempt Inspector is a free application that can be used to scan the enterprise for weak passwords by using dictionaries of compromised or weak credentials, as well as password hashes from past known breaches. The tool can also analyse current password policy, as well as scan for 'stealthy' admins that have administrative access but are not managed via a security group in Active Directory.

See https://www.preempt.com/preempt-inspector

With 4 out of 5 breaches starting with credential threat, "identity is the perimeter".

CyCognito

Founded	2017 Headquarters: Palo Alto, CA
Founder(s)	Rob N. Gurzeev
Investment Funding	Early stage, seed start-up in stealth
Website	www.cycognito.com
Sector	Cyber, Infrastructure, Networking
Key Points	• Cloud-based network security analysis platform designed to discover organisations' security weak spots • External attack simulation platform • No installation or corporate resources required

Overview

CyCognito is still operating in "stealth" mode, having completed an incubator and early accelerator programme with UpWest Labs. They perform deep scans across organisation's externally facing infrastructure, not just its firewalls, for vulnerabilities and changes that could be exploited and report upon the potential impact if they were to be exploited.

CyCognito's platform provides organisations a complete Attack Surface analysis from a sophisticated attacker point of view. By understanding which specific Attack Vectors can lead to a data breach, security analysts and management can eliminate their cybersecurity threats effectively.

Technology

Sophisticated attackers seek the Path of Least Resistance to compromise an organisation's network, not necessarily most critical CVE or asset. Mitigating these potential Attack Vectors requires continuous and comprehensive Attack Surface Analysis to identify the most attractive security weak spots in the eyes of the attacker. CyCognito's SaaS platform continuously simulates sophisticated attackers' actual reconnaissance and examination processes across live infrastructure and network assets to provide comprehensive Attack Surface Analysis in real-time.

CyCognito enables security teams to eliminate critical attack vectors in:

- known and unknown assets
- internal and 3rd party components
- on-premise and cloud environments
- software vulnerabilities and implementation flaws
- infrastructure and applications

What an attacker can see, CyCognito enables security teams to see and remediate in advance. This is achieved through a world-wide, bot-like network that indexes organisations and network assets to produce a graph of what's owned or closely related (similar to a Google index).

Applications

Working for intelligence organisations and assisting them in establishing new infrastructure for offensive security, CyCognito founders learned that for attackers, the road to glory is the path of least resistance. Unlike penetration testers and security researchers, attackers do not seek medals or bonuses for solving complex challenges. This is true for both state-level actors and individual cybercriminals. Their sole objective is to act in a cost-effective, stealthy manner in their pursuit of information or money.

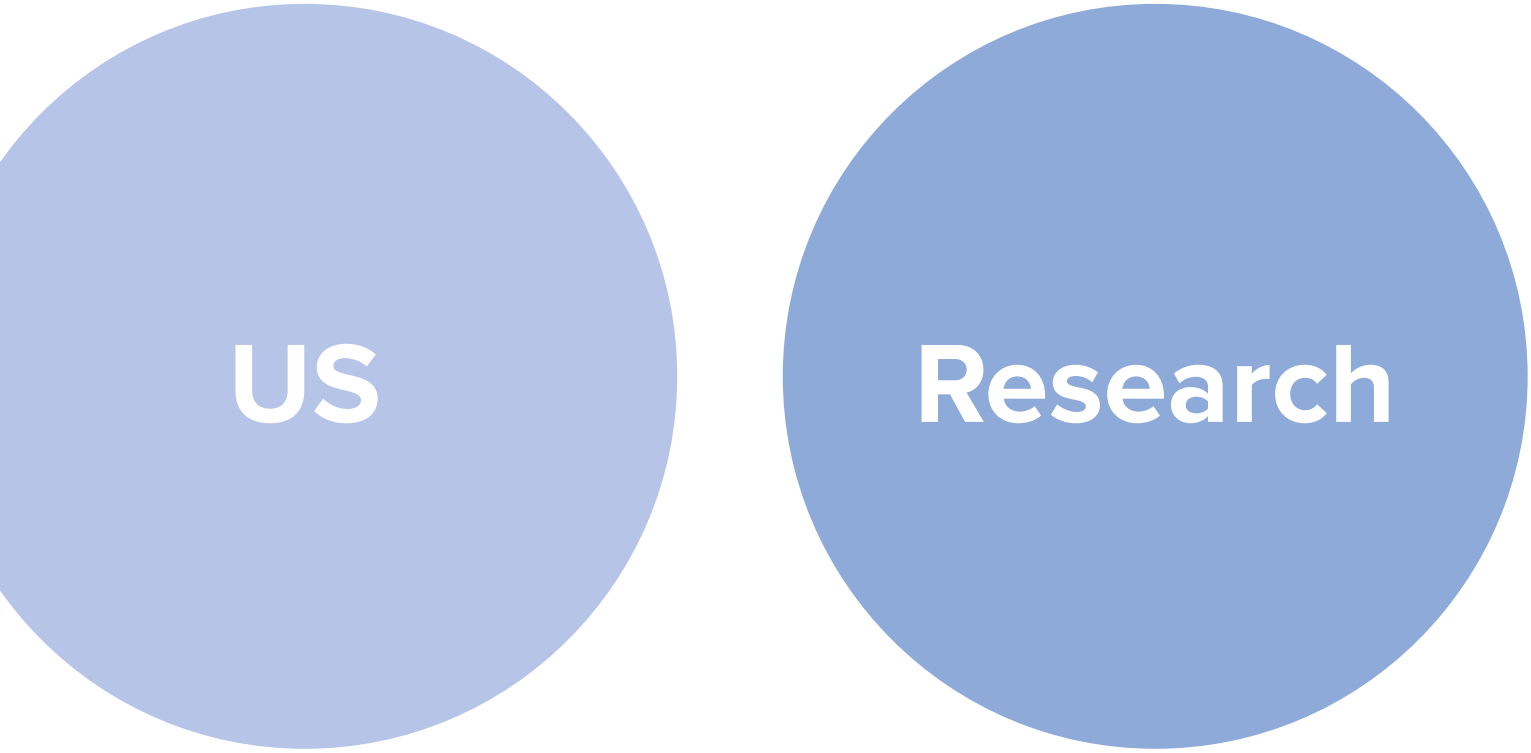
Organisations must invest significant resources in trying to understand how attackers see their attack surface and what they can actually exploit, as opposed to scanning the assets they already know for security issues, that even if found are often of minimal or no interest to attackers. This could be seen as a mindset shift from a discussion focused on CVE's and vulnerability score, to one focusing on attack vectors discoverability, attractiveness and exploitability. This can be executed only by an external actor or system that receives no prior input regarding the target network or cooperation from the organisation.

With CyCognito, clients need only provide their company name. From this, they can map the entire attack surface and examine what exploits exist.

Observations

High quality penetration testing is very expensive, and every change within an organisation's network (new applications, servers, configurations, etc.) requires a new test, practically starting from scratch. It's not scalable at all. Therefore, only a product efficiently incorporating a black box approach in an automatic and scalable fashion could become a game changer in the inherently asymmetrical race between attackers and defenders.

Friday 25 May



Tour

Ntegra Greenside

Knowledge for IT Leaders

Sequoia Capital

Founded	1972 Headquarters: Menlo Park, California
Founder(s)	Don Valentine
Investment Funding	Venture Capital: Early Stage Venture, Late Stage Venture, Seed 277 Exits
Website	www.sequoiacap.com
Sector	Angel Investment, Finance, Impact Investing, Venture Capital
Key Points	• VC firm focused on energy, financial, enterprise, healthcare, internet, and mobile start-ups

Overview

Sequoia Capital is a venture capital firm founded in 1972. The Wall Street Journal called Sequoia Capital "one of the highest calibre and most influential venture-capital firms". It primarily focuses on start-up/seed, early stage and expansion investments. Sequoia has funded an unprecedented number of enormously successful companies including Google, Yahoo, PayPal, Electronic Arts, NVIDIA, Cisco Systems, Oracle, Apple and YouTube. In aggregate Sequoia backed companies account for more than 20% of NASDAQ's total value.

Companies presented

- Orbital Insights
- Big Panda
- ThousandEyes
- Snowflake
- Front

Summary

Jeff Pashalidies kicked off the day and closed the tour with an overview of future trends that Sequoia is tracking.

Orbital Insights, Inc.

Founded	2013 Headquarters: Mountain View, California
Founder(s)	James Crawford
Investment Funding	Total Funding: $78.7M Last Funding: Series C
Website	www.orbitalinsight.com
Sector	Analytics, Big Data, Financial Services, Geospatial, Software
Key Points	• Provides a "macroscope", finding truth and transparency in the ever-expanding supply of satellite images • A Geospatial Big Data company leveraging rapidly growing availability of satellite, UAV, and other geospatial data sources • The company's goal is to understand and characterise socio-economic trends at global, regional, and hyperlocal scales

Overview

Orbital Insight combines the wealth of publicly available satellite imagery with image processing, machine learning and analytics to provide unique information and insights. This can provide insight into movements of people and machines, installations, geological changes, construction, illumination and anything that can be 'seen from above'. We had previously seen Orbital during the 2016 tour, but Sequoia considered their advances appropriate to showcase them again.

Technology

Orbital Insight provides insight based on imagery then analytics using a combination of AI, machine learning and machine-human collaboration.

The capability is built in an Amazon Web Services (AWS) cloud and makes use of several Elastic Compute (EC2) instances running both Computer Vision (CV) algorithms to process images and Data Science algorithms to process numbers and/or volumes. The combined output from these delivers usable insight.

It also uses "deep learning" techniques (looking for anomalies using machine learning / neural networks) to identify unknown unknowns.

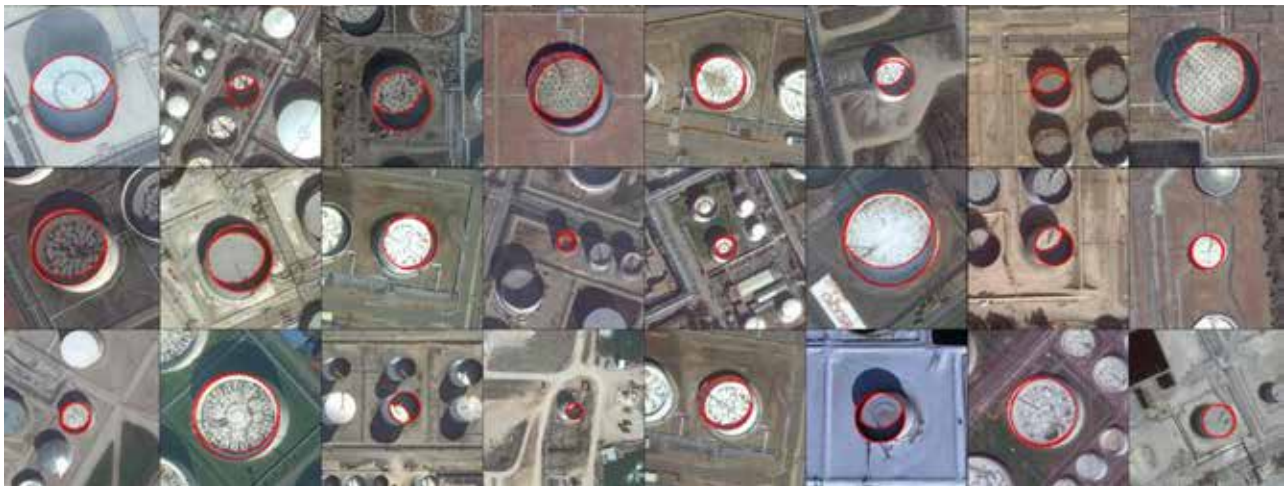
Applications

Insights often start out as a search for the patterns of life such as "What does normal look like?" This can be analysed to look for anomalies and predictive analytics can be used to accurately predict future states. Also, this can start out as a wide area search where it is not clear what is being looked at or what users are trying to monitor. This can then result in new insights based on following an individuals' lines of interest. Orbital Insights have combined AI, Cloud and GPU capabilities to create a Geospatial AI company for the commercialisation of Space; iTunes for satellites.

Orbital Insight most often work by partnering with organisations who have ideas, theories or initial analysis that can be proved or enhanced by the analytics and insights from earth imagery. Examples of this include:

- car counting at Walmart car park - this can be used as a proxy for people going to the store and so gives both current figures and can predict footfall. Used by analysts, this information can also give some insight into individual store performance which, when rolled up can be compared to retail predictions to spot gaps and invest in stock accordingly.
- identification of oil storage facilities - looking for the distinctive shape of oil storage tanks can easily be automated. The results can be analysed by assessing the level in the tanks to provide an accurate estimate of the total storage capability and the actual amount stored. This can provide an accurate view of the strategic reserves.
- predicting the rate of economic growth - this can be done through identifying and assessing changing levels of construction, illumination, car usage, pollution and others.

If it can be seen from above (buildings, trucks, ships, trains, aeroplanes, oil storage, cars, housing, etc.) Orbital Insights can combine geo-located images with business and other intelligence.



Observations

Big Data and Predictive Analytics have been discussed and hyped for several years as part of the big four IT topics i.e. cloud, mobile, social and analytics. However analytics is only as good as the data you can provide to analyse and so a new source of data is always going to provide the capability for new and interesting insights. The use of satellite imagery combined with powerful image processing and analytics tools has the capability to provide highly relevant insights on a global scale.

Over 260,000 US car parks are overflown and imaged every day. Improvements in electro-optical combined with synthetic aperture radar (that enables “vision” through clouds) as well as improving geo-spatial, terrestrial sources are enabling Orbital to see trends first.

Their platform is currently able to ingest 10x more than it is currently using to provide commercial and defence insights, and their ML capabilities are achieving 96-97% precision with concerted training (where as human intelligence agents are achieving 95% at best), enabling Orbital to develop autonomous anomaly detection and semantic change detection algorithms to seamlessly integrate with digital operations.

BigPanda

Founded	2012 Headquarters: Mountain View, California
Founder(s)	Assaf Resnick, Elik Eizenberg
Investment Funding	Total Funding: $51M Last Funding: Series B
Website	bigpanda.io
Sector	Cloud Infrastructure, Developer Tools, Enterprise Software, Infrastructure, IT Management, SaaS
Key Points	• Enables large enterprises to automate and scale IT Operations to meet the demands of digital transformation • Provides an Autonomous Digital Operations platform for IT incident management • Backed by top-tier investors including Sequoia Capital, Mayfield, and Battery Ventures

Overview

Big Panda's Algorithmic Service Operations platform is the first to achieve Service Organisation Control type II (SOC2) certification in data security, availability and confidentiality. SOC2 is a widely recognised standard issued by the American Institute of Certified Public Accountants (AICPA) which measures a standardised set of security and data practice criteria, requirements and controls. Companies including publicly traded enterprises, financial firms and healthcare organisations have compliance requirements that require SOC2 audits.

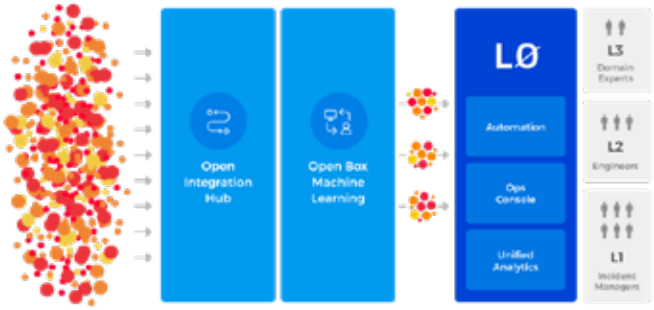
Business is counting on Ops to keep critical services and applications running smoothly 24×7. But ever-increasing complexity in infrastructure means teams are in constant firefighting mode. The result, services are frequently down, SLAs are repeatedly violated, and customers are increasingly frustrated. It is a vicious cycle that IT Ops team just can't seem to escape. This scenario, repeated daily in NOCs and Ops Centres inside large, complex enterprises, takes a toll on business's ability to innovate and execute on critical digital transformation initiatives.

It doesn't have to be like this. Autonomous Digital Operations (ADO) can help.

Technology

Autonomous Digital Operations is the next generation of IT automation. ADO moves enterprise IT Ops from legacy, rules-based solutions to a fully autonomous future. ADO eases the burden on IT Ops by dramatically improving time to resolution, increasing service availability, and driving down operational costs. ADO ultimately helps IT teams to support and accelerate digital transformation initiatives.

IT incident management spans five discrete stages: detection; prioritisation; investigation; remediation; and escalation. ADO uses advanced Machine Learning to deliver autonomous responses to IT incidents throughout this lifecycle. For some incident types, ADO can handle all incident management stages autonomously. Incidents that can't be autonomously resolved are escalated to Level 1 teams for resolution. AOD can also be configured to bypass



Level 1 teams and escalate certain incident types directly to Level 2 or Level 3 teams based on organisational structure and workflows.

Each enterprise and IT Ops team knows the best way to customise and adopt technology to match its unique needs. In the context of ADO, it is important that each team can adopt ADO at its own pace and choose different levels of autonomy for different needs of its organization. Therefore, ADO platforms allow IT Ops teams to configure the level of automation (M/L assisted vs. fully autonomous) that works best for them. ADO platforms enable this granular control at the organisational, application and service levels.

Like autonomous cars that have evolved from manual to automation-assist to a path to fully autonomous driving, ADO is leveraging advanced Machine Learning techniques in its rapid evolution. Today ADO Platforms perform autonomous incident management where possible and request human intervention when required. In future, ADO Platforms will perform autonomous incident management under most conditions and request human intervention only under extreme circumstances. Beyond that, ADO platforms will mature to perform autonomous incident management under all conditions.

Applications

Autonomous Digital Operations (ADO) Platforms have three tenets

- they are based on Machine Learning
- they provide an Autonomous Layer that automates incident management lifecycle
- they provide an Integration Hub to tie together all existing IT tools

ADO's Machine Learning engine delivers real-time automation logic that continuously learns from and adapts to dynamic IT environments. The core tenets of Machine Learning (ML) include:

- Transparency: Ops teams can easily visualise and understand the automation logic created by the ML engine
- Trust: The ML engine is highly deterministic and generates the same results reliably time after time
- Control: Ops teams can augment the ML automation logic with their situational knowledge that is unique to your organisation

ADO's L0 Autonomous Layer leverages machine learning to automate significant portions of the incident management lifecycle. What can't be automated is seamlessly escalated to Ops teams for resolution. The core tenets of the L0 Autonomous Layer include:

- Transparency: Ops teams can easily visualise and understand the automation logic used by L0 to automate all aspects of incident response
- Configurability: Ops teams have complete control over the level of automation (M/L assisted vs. fully autonomous) at a highly granular level (e.g. per team, application, incident type, etc.)
- Control: L0 provides Ops teams complete control regarding which teams (e.g. L1, L2, L3 teams) to escalate incidents to, when, and how

ADO's Integration Hub seamlessly connects with a diverse and ever-changing ecosystem of legacy and modern tools across different technology domains.

The core tenets of the Integration Hub include:

- Time-to-Value: Within weeks, Ops teams can easily integrate all six categories of tools (monitoring, topology, change management, collaboration, business intelligence and remediation)
- Real-time Insights: The hub supports real-time high-speed data ingest and integration across these tools silos to unlock and share powerful insights
- Future-Proof: Ops teams can easily absorb and integrate new tools stacks that organisations adopt daily

Observations

Operational scale and complexity is an everyday fact at large, dynamic enterprises. Most IT tools are underutilised or under-implemented. "Tool-sprawl" plagues many organisation, which in turn encourages data silos that lead to confusion. It's clear that today's Infrastructure and Operations teams are drowning in too much data, but not enough insight. When it comes to Infrastructure and Operations' role in digital transformation, focus on improving business outcomes over lowering IT operational costs.

Some organisations seem almost culturally resistant to automation. But Infrastructure and Operations leaders are actively seeking solutions to help them better manage the chaos. Failure to adopt artificial intelligence and ML should no longer be an option and organisations need to start working on deploying machine learning now!

Automation is only the beginning of the digital transformation journey, not the result. Enterprises must stop implementing automation in an opportunistic, ad hoc fashion and adopt a systematic, end-to-end automation approach.

ThousandEyes

Founded	2010 Headquarters: San Francisco, California
Founder(s)	Mohit Lad, Ricardo Oliveira
Investment Funding	Total Funding: $60.7M Last Funding: Series C
Website	www.thousandeyes.com
Sector	Customer Service, Enterprise Software, Network Hardware
Key Points	• Network Intelligence platform that delivers visibility into every network modern enterprises rely on • Establishes clear baselines, success criteria, operational metrics and proactive problem remediation

Overview

ThousandEyes is a network intelligence platform that provides insights across the networks and connected devices that an organisation depends upon. It provides visibility to network engineers, operations teams, support/service desk staff, developers, and business users with an accurate, up-to-the-moment understanding of what's happening in an organisations internal and external networks.

ThousandEyes were first presented to the Tour by Sequoia three years ago in 2015. The company was founded on the understanding that the Internet was going to impact people, companies and networks in profound ways. The two founders were researching Internet routing and network protocols at UCLA when they founded the company in 2010 to "empower companies and people to see, understand and improve connected experiences everywhere". Today they have offices in San Francisco, Austin, New York, Dublin, London and Tokyo.

Technology

We've seen a lot of corporate IT teams struggling with a big change. All apps and services lived in the data centre, all users were in company branches and connected by MPLS networks that were predictable and had a Service Level Agreement. Monitoring was based on passive data from their own devices: traffic flow, packet capture and device data, plus APM for apps and code they controlled.

The world has changed dramatically. Business relies on a large and growing set of external dependencies that it doesn't own or control. For customer digital experience, that means DNS, CDN, ISP and DDoS security providers. For SaaS and cloud adoption that means ISP, secure web gateway, public cloud and SaaS providers. This means that every communication cross paths that include multiple network organisations through the Internet. The Internet itself is an unpredictable factor with BGP routing changing paths constantly.

Passive network data is useful but can't be retrieved from all those external places. It's not possible to gather packet captures from the CDN. You can't use SNMP on devices in your ISP. You can't do APM code injection into Microsoft 365. This creates operational blindness and it's not just a pain, it's a material impact for the business.

Operational blindness matters because the cloud creates new operational processes. For IT infrastructure and networks, software companies control highly trained engineers who can correlate different data sets then directly fix issues. For cloud and external providers, companies need enough evidence to first identify which

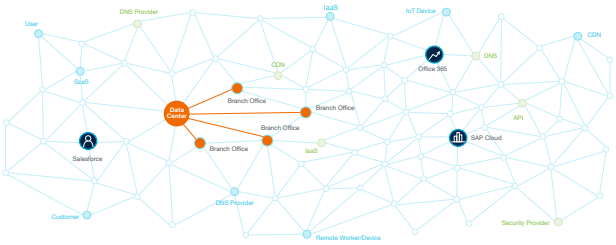
provider or network domain the problem lies in and then need enough evidence to persuade an external provider to act. Sometimes that provider must in turn get one of their providers to act...

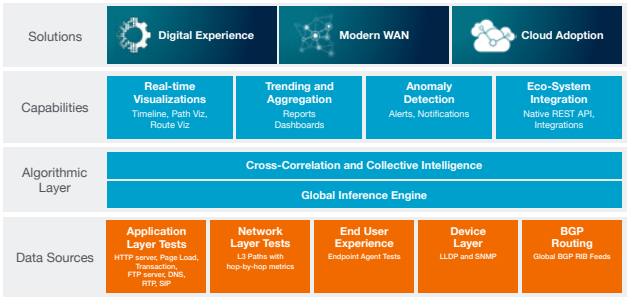
Without evidence at multiple layers, both app and network can't overcome plausible deniability and escalate effectively. Mean time to recovery not only skyrockets but in some cases resolution time is unknowable.

ThousandEyes was created to deal with these issues and to help companies understand the experience gained from every user, to every app, over any network, especially the Internet. They have built a cloud-based platform and app that allows them to collect and synthesise data from many different vantage points and created three types of smart monitoring agents to cover every thread of internal and external communication that matters to a business:

- Cloud Agent - globally distributed agents installed and managed by ThousandEyes in 150+ POPs around the world
- Enterprise Agent - lightweight software-based agents, easily installed on enterprise networks, in data centres, branch offices & VPCs
- Endpoint Agent - browser-based plugins installed on end-user laptops and desktops

The breadth and depth of data sources are from app, to network, to network devices, to Internet routing layers. Add in real user experience data, Big Data-powered global inference engine learns and improves the data model accuracy over time. ThousandEyes add 'collective intelligence' such as Internet outages. Cross correlation algorithms turn all this data into intuitive visuals and easily shareable reports and dashboards. They automate via alerts and notifications, plus bi-directional integration through native REST APIs.





Applications

ThousandEyes customers take advantage of these capabilities to address three major use cases areas: improving digital experience; WAN modernisation; and cloud adoption. The data created and collected by the agents makes it possible to see every network as if it was your own. Customers gain a categorical understanding of app and service delivery across the Internet, through the enterprise, and to the endpoint. They can correlate app and service performance with every underlying dependency from DNS to CDN and security providers, across end-to-end network paths and BGP routing and go deep into QoS remarking, PTMU discovery, MPLS tunnels and Wi-Fi interference to get to the root cause of an issue, quickly.

Observations

ThousandEyes provides visibility into and between customer datacentres, branch offices, SaaS, public cloud providers, customers and remote workers. Their unique global dataset gives access to new insights. They harness the collective intelligence of tens of thousands of concurrent monitoring tests across the globe and identify Internet outages and reachability problems and show them directly in context of your app and service delivery, leading to is dramatically lower time to recovery.

ThousandEyes visibility establishes clear baselines, success criteria, operational metrics and proactive problem remediation, leading to more successful deployments and cost-efficient operations that deliver better user experiences.

Snowflake Computing

Founded	2012 Headquarters: San Mateo, California
Founder(s)	Benoit Dageville, Marcin Zukowski, Thierry Cruanes
Investment Funding	Total Funding: $472.9M Last Funding: Series E
Website	www.snowflake.net
Sector	Analytics, Cloud Data Services, SaaS
Key Points	• Data warehouse built for the cloud • Combines the power of data warehousing, the flexibility of big data platforms and the elasticity of the cloud

Overview

Positioned as a Challenger in Gartner's 2018 Magic Quadrant for Data Management Solutions for Analytics (DMSA). On the one hand there's a lot of hype about new data types, new technologies to store and manage them efficiently and new roles and skills to use them effectively; on the other, there's recognition that investment in foundational traditional technologies will be essential to serve as a platform for the next wave of innovation. Snowflake supports both sides of the conversation.

Snowflake was founded by a team of database experts and industry veterans from companies including Oracle, Teradata, Actian, and Cloudera. They have significant backing from established investors who have a track record of success, ensuring that they have the resources to continue to invest in their product and growth. They've been adopted by a rapidly growing number of customers since first use of the product in 2014, and ~1000 customers have signed up for Snowflake's 'Data Warehouse Built for the Cloud' to date. Snowflake's mission is to enable every organisation to be data driven.

Technology

Systems built in the 80s, 90s and 2000s were conceived before the cloud. They weren't designed to operate at the performance, scale and simplicity required today. Snowflake has been built for the cloud, to leverage the capabilities of the cloud, making it very different from anything else in the market. Snowflake provides:

- a complete analytic relational database built for the cloud
- SQL compatibility with Oracle, Teradata, Netezza, SQL Server, Vertica
- a completely self-managing solution that eliminates all operational expense
- essentially unlimited data sizes
- essentially unlimited concurrency
- support for both structured and machine generated, semi-structured data
- a much lower cost base than traditional SQL data warehouses

Snowflake removes the constraints that prevent businesses from working with all their data. They're focused on making Snowflake as fast as possible and effectively scalable to an infinite number of concurrent queries against

the same data set, with no limits on the number of users. Older systems like Oracle and more recent open-source systems such as Hadoop were built with the view that everything had to be finely tuned by administrators. Instead, Snowflake doesn't require any manual tuning and is optimised across a wide variety of queries.

An AWS region is the primary scale unit for Snowflake. Within an AWS region, a single Snowflake VPC can support many thousands of multi-tenant customers. VPS customers are provisioned with a dedicated VPC. All servers which instantiate the customer encryption keys as well as the metadata store are fully dedicated to the customer. All communications between servers are fully encrypted.

VPS customers are provisioned with their own Amazon account for S3 storage. The HSM is configured with a separate partition that is dedicated to the customer. All data stored within S3 is fully encrypted.

Snowflake configures and maintains the dedicated customer VPC by connecting this using a VPC bridge to their production multi-tenant VPC. Scale compute up and down, transparently and automatically. Data storage is columnar compressed.

Being a complete SQL database means Snowflake works with the tools business users understand such as ETL and analytics. Snowflake is a fully ACID-compliant relational database that supports role-based security, multi-statement transactions and a host of other features.

Snowfalke's architecture supports all data types:

- Structured and semi-structured data (JSON, XML, Avro)
- Centralised storage of data, accessible by any user & application
- Multi-petabyte scale
- Query petabytes of data in as little as a few seconds

Applications

CapitalOne have a simple charter: they want to give a personalised and engrossing experience to all their users. Unfortunately, their technology before Snowflake did not support that. Snowflake is up to 200x faster than solutions not built for the cloud. It maintains a consistent SLA as resources grow and shrink automatically, loading data does not impact query performance. Multiple groups access data at the same time with no performance degradation. It supports an unlimited number of simultaneous users.

With compute truly separated from storage, Snowflake effectively allows an infinite number of compute clusters to operate on the same data at the same time.

Observations

Snowflake is a fully elastic cloud data warehouse in which compute is separate from storage, so you pay for things separately. Storage is now priced at $30/terabyte for a one-year commitment, meaning you can effectively store data in the world's most powerful data warehouse for the same price as Amazon S3.

With Snowflake you can also share data both internally and externally with business partners, with no data movement. Data consumers directly access the shared data, the data doesn't need to be moved or copied and hence, they always see the current version of data, even when the provider updates data while it is being consumed. Consumers can immediately start querying shared data and combining it with their own data by just launch a virtual warehouse on-the-fly.

Front

Founded	2013 Headquarters: San Francisco, California
Founder(s)	Laurent Perrin, Mathilde Collin, Thibaud Elziere
Investment Funding	Total Funding: $79.3M
Website	frontapp.com
Sector	Apps, Collaboration, Email, SaaS, Software, Unified Communications
Key Points	• Shared inbox for teams • Aims to change the way teams get work done • Messaged from email, Texts, Slack and social media, all in one place

Overview

Front redefines work communication with the first shared inbox for teams. By unifying email, customer communication channels and apps in one platform, Front helps teams collaborate efficiently, have more context and visibility into every conversation and to work faster and better together.

This year Front, went on a five-day fundraising sprint to raise a series B round. They pitched to 11 investors, received 10 investment offers and walked away with $66 million from Sequoia Capital and others. Look at their Series B pitch deck at:

https://medium.com/@collinmathilde/front-series-b-deck-6dc686267a24

Technology

How work gets done hasn't changed much for many years. Email protocol is not built for collaboration and was not designed for teams. Email cc's, bcc's and forwards slow down productivity, confusing accountability and ownership. Front makes email collaborative allowing users to collaborate on a single email, increasing transparency and clarifying accountability and ownership. It centralises collaboration for more efficient and effective customer communications bringing all emails, apps and teammates together in one place.

Front is hosted on AWS and provides workflow and a collaboration chat feature that is stored in a Front Server, enabling users to collaborate with their team "behind-the-scenes" without generating even more email.

Front features:

- Assignments - Choose an assignee for clear ownership of every conversation so no message slips through the cracks
- Internal-only comments - Coordinate next steps with in-line comments and @mentions, without the unnecessary additional emails
- Collision detection - See when teammates are typing in real-time to work together on responses and avoid duplicate replies
- Rules - Work faster and act on messages automatically using tags and keywords to trigger rule-based workflows

- Reminders - Follow-up is critical. Set custom reminders to help stay on top of every conversation
- Canned responses - Answer common questions with one click and save time with message templates to jump start replies
- Analytics and reporting - Measure performance metrics like response time and message volume for more data-driven decisions
- Open API - Customise Front with a flexible API that provides complete data access for custom integrations
- Integrations - Create and update Salesforce records, JIRA tasks and more to keep all apps in sync from the inbox

IMAP sync for Gmail and Office 365 make it easy to get started with Front. Adding other channels is simple with SMTP redirects and OAuth for other messaging services. Admin controls like audit logs and access permissions keep data secure in Front. SAML Single Sign-On and two-factor authentication secure team logins safely. Front provides the flexibility and admin controls needed to use it across the company and enables Easy management of multiple, separate user groups with a single Front account.

Applications

For consulting, real estate, accounting and law firms, clients are their business. Front enables services organisations to build an amazing client experience by connecting all the tools they rely on within their inbox. Route messages to the right team and enable them to respond to clients with context and speed.

Observations

Front proves that you don't have to sacrifice ease-of-use for efficiency.

Closing statements

Ntegra would like to thank all the Greenside US Research Tour 2018 stakeholders for their support and participation in a hugely successful event which we thoroughly enjoyed. In closing this year's report, we'd like to add the following thoughts:

The technological revolution is becoming as disruptive as the industrial revolution, bringing amazing opportunities but also new forms inequalities. If we want technological change to benefit everyone, we need to start preparing now. Most people are optimistic about their own ability to adapt, with workers saying they are confident they can update their skills if automation and algorithms impact their jobs. Everyone should be able to take advantage of the positive opportunities that technology brings, e.g. supporting people who want to work more flexibly or helping people to move from routine, repetitive tasks into more rewarding work. Harnessing technological change for good also requires vigilance and action to prevent new technology being used to control rather than liberate the workforce.

Throughout history the technology of warfare has been primarily focused on defence. Well-constructed castles with thick stone walls were once considered unassailable. However, with the advent of nuclear weapons, our defensive systems have become virtually hopeless. Underground bunkers, caches of food and identifying safe water sources are pitiful defences against the destructive power of a direct hit.

We have seen something similar with cyber-defence where, until recently, building safer, stronger, more secure 'castles' has been the objective and main driver of innovation. This year felt slightly different with several of the presenting companies adopting more of an aggressor's role; taking the fight to the cyber criminals. There were no mentions of the 'cyber kill chain', a paradigm in which it is assumed we're already under attack, but a more proactive stance to stopping attacks before they happen.

Innovation watchers in Silicon Valley and further afield are split between anticipating the next billion scale platform S-curve and those looking for new opportunities to emerge, supported by maturing technologies. Whichever way things go, organisations must keep-up and be prepared for "the next big thing". In this spirit, the Ntegra Greenside Programme is introducing new services to help organisations systematise their own internal innovation processes and provide an agile Proving Environment where new technologies and operating models can be spun-up and assessed.

Organisations should think digital but act analogue. That is, they must remember the purpose of innovation is happy people. Self-driving, digital, trucks may seem like a cool idea but truckdrivers may not be so happy.

We hope you enjoyed this year's tour and are looking forward as much as we are to whatever 2019 brings.

Company Index

Ntegra Greenside

Knowledge for IT Leaders