(https://www.123rf.com/photo_27317373_stock-vector-vector-circle-based-owl-head-symbol-set.html)

[logo, name, and acronym TBC]

Decentralized Video Analytics (DEVA)

An open, decentralized and scalable blockchain-enabled video A.I. platform

Whitepaper

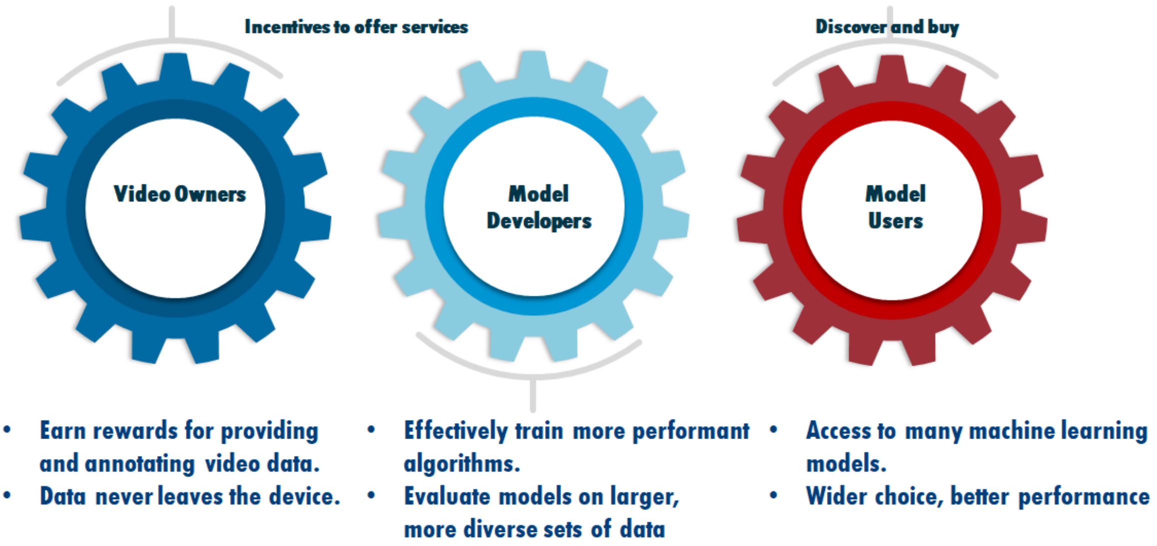
v0.6, November 2018

[Insert ToC]

Introduction

Video analytics is a rapidly growing multi-billion-dollar market. Yet, developing video analytics applications is currently highly challenging, as developers must acquire a large amount of raw video data, process and annotate the videos, design effective machine learning models, and train/tune the models on high-performance servers. To do these requires a variety of resources, including data, domain-specific knowledge (for video annotation), machine learning expertise, and computation resources.

Our Decentralized Video Analytics (DEVA) platform aims to create an open, decentralized, and scalable video analytics ecosystem, which enables a multitude of participants – including companies, government agencies, and individuals – to act as active contributors and/or users, while at the same time providing economic mechanisms and incentives for making this viable and sustainable.



Video sources, analysis methods, the resulting analytics metadata, and related services thus are commoditized to become economic goods that can be supplied, traded, and valued by all participants. We foresee such a disruptive infrastructure to foster rapid innovation in video analytics and generate incentives for innovative offerings.

The key novel aspects of this infrastructure include:

- Video analytics for multiple use cases;
- Decentralized data acquisition and annotation;
- Federated machine learning;
- Privacy-preserving model training;
- Blockchain-based contracts and payments;
- Market mechanisms to facilitate economic transactions through tokens;
- Market mechanisms to determine transaction prices.

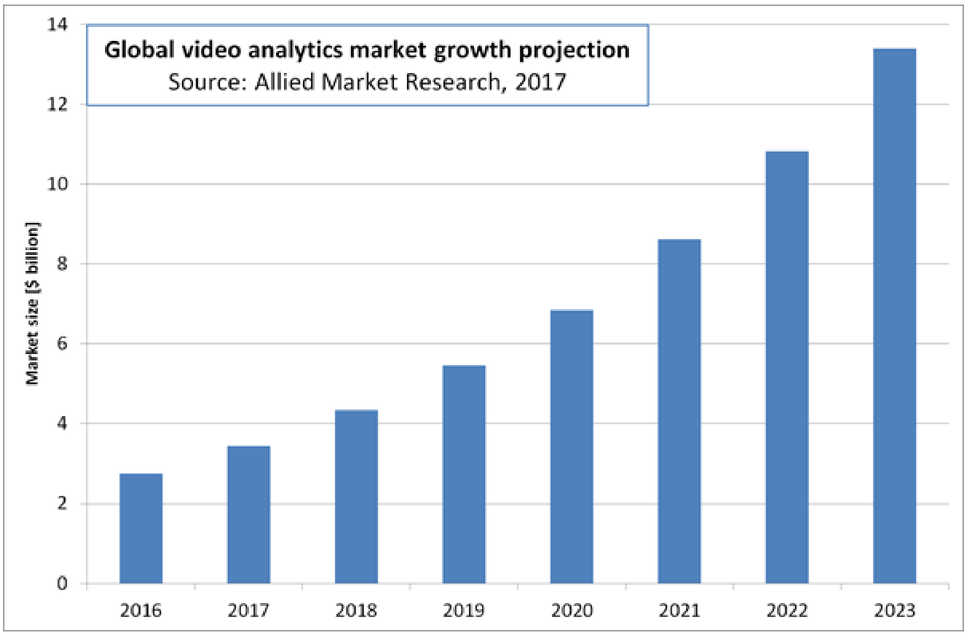
DEVA thus represents a paradigm shift in the way video analytics are done today. Today's closed vertical systems from a single provider lead to inefficient duplication of efforts and fail to architecturally address the needs for flexibility and wide participation; existing video services suffer from the same shortcomings. Only an open video architecture supporting a federated ecosystem that retains individual user control yet enables and incentivizes sharing will achieve widespread acceptance and usage. It is also the first blockchain platform with support for Trusted Execution Environment (TEE) technologies to ascertain confidentiality and privacy.

Video Analytics Market

According to IHS estimates, there were 245 million professionally installed video surveillance cameras active and operational globally in 2014. Asia accounted for 65 percent of installed security cameras. Aside from CCTV cameras, cameras are now ubiquitous: mobile phones, in-car cameras, drones are equipped with high-quality cameras that are utilized to capture a huge amount of video data.

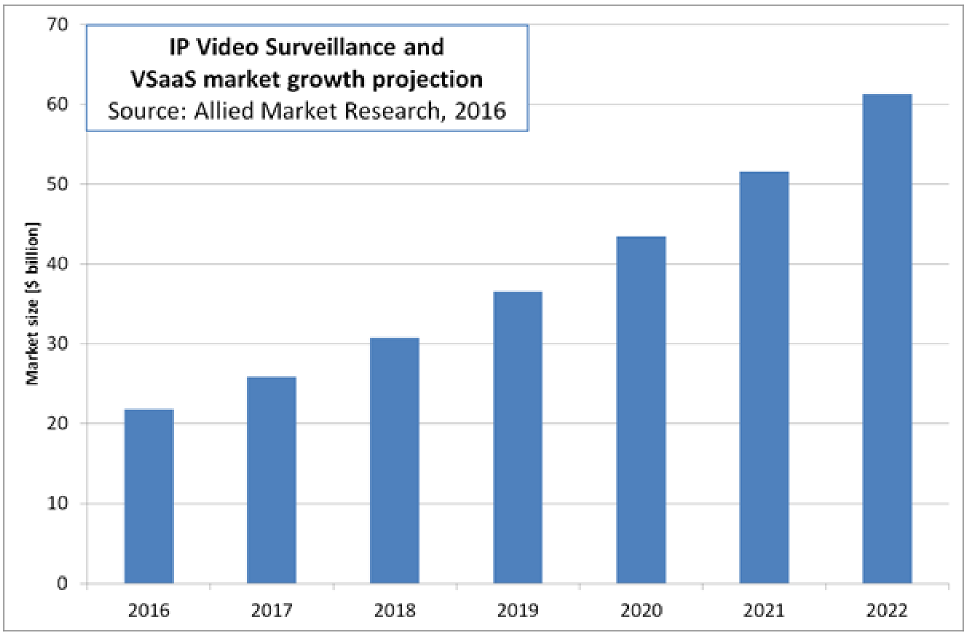
The number of cameras installed in private homes, shops, restaurants, gas stations, and offices is growing every year and so is the demand for more advanced solutions. Increasing demand for video analytics in smart devices, such as smartphones, laptops, tablets, personal digital assistants, and drones, which are used for both personal and business purposes, presents further growth opportunities in this market.

The global video analytics market is expected to witness significant growth rate, due to the rise in smart cities expenditure by governments of various developed and developing regions, including North America, Europe, and Asia-Pacific. According to an Allied Market Research report, the global video analytics market was valued at $2.7 billion in 2016, and is estimated to reach $13.4 billion by 2023, growing at a CAGR of 25.7%. The demand for intelligent video analytics systems and services capable of analyzing video streams in real time has grown significantly and will continue to grow.



Today video surveillance is an indispensable element of security systems, both in private and public areas. It is one of the components of intelligent transport systems and an indispensable element of smart cities' infrastructure, which is one of the fastest growing markets. In smart cities, CCTV is used to record violations of traffic rules, to control the operation of municipal services, to restrict access to closed areas, as well as in parking and toll payment systems on high-speed roads.

According to an Allied Market Research report, the global IP video surveillance and Video Surveillance as a Service (VSaaS) market was valued at $21.8 billion in 2016, and is estimated to reach $61.3 billion by 2022, growing at a CAGR of 18.8%. IP video surveillance has revolutionized the video surveillance industry by reducing costs and offering a high level of flexibility and scalability in comparison to conventional systems. IP cameras allow their users to employ innovative technologies for video analytics.



Video analytics provides functionalities such as object detection, object identification, object recognition, object counting, event detection, event recognition, anomaly detection, and many others. “Objects” in this context can be people, faces, animals, vehicles, items, etc.

There is an increase in the adoption of cloud, computing services, connected devices, and video surveillance by governments to enable smart city functionality and combat crime & terrorism. The need for situation awareness and behavior analysis is expected to drive the video analytics market in the future.

The latest developments in the fields of deep learning and machine vision have allowed video analytics algorithms to reach new levels of accuracy. Technologies powered by deep neural networks and artificial intelligence permit the analysis of video at superhuman levels. It is widely accepted that video analytics is one of the best applications of AI using deep learning.

Ecosystem

Developing video analytics applications is currently highly challenging, as the app developer must acquire a large amount of raw video data, process and annotate the videos, design effective machine learning models, and train/tune the models on high-performance servers. To do these requires a variety of resources, including data, domain-specific knowledge (for video annotation), machine learning expertise, and computation resources. Existing platforms only help to some extent, for example, data and annotations could be collected through crowdsourcing, model design could be outsourced via competitions, and model training could be done on the cloud. However, there is no integrated process combining these platforms. Further, most platforms nowadays are centralized, such as Amazon Mechanical Turk for crowdsourcing, Kaggle for model building competitions, and Amazon EC2 for cloud computing. This poses additional confidentiality and privacy issues. The latter is increasingly recognized around the world, with governments tightening regulations and organizations becoming increasingly careful in offering their data to third parties.

DEVA provides huge benefits for small- and medium-sized enterprises (SMEs) or individuals, who do not have access to the data they require for their analytics problems, and who lack the resources and/or know-how to build customized solutions and may not be able to afford the necessary computing infrastructure. However, even large companies and government agencies can benefit from DEVA and save costs because of the sharing of data sources, algorithms, and implementations, which reduces duplication of efforts and reduces the need for expensive customized solutions. We next discuss the issues DEVA aims to solve through a number of concrete scenarios.

Scenario 1: Autonomous Vehicles

A company is developing an in-car driving assistant application (such as Tesla's Autopilot) based on video feeds from cameras. The application needs video analytics models to recognize dangerous situations, such as sudden overtaking (possible from the right side), or a careless bicycle rider during nighttime.

To build such models, the driving assistant maker would have to hire human drivers to drive around, record each driver's views and moves, label potentially dangerous situations manually, and finally build a video analytics model – an expensive and time-consuming process. Alternatively, the company could crowdsource dash camera videos and annotations, and publish a prized competition for model building. However, this would mean that the purchased data will be revealed to the competition platform as well as all participants of the competition, which may be neither desirable nor legally possible.

Note that incentives are not well aligned in the above workflow: a data owner does not earn much by selling raw data; a highly skilled model builder focuses on winning the competition, rather than creating practical models; a moderately-skilled model builder is not strongly motivated as she has little chance to win the prize. Hence such a centrally controlled workflow (as in the previous example) is too restrictive by dictating the roles and activities for the rest of the contributors. But this is not necessary as we discuss next.

Our solution to the above problem is DEVA, a decentralized, integrated and flexible data value chain ecosystem powered by blockchain technology.

In the above scenario, the application developer only needs to define the problem (such as providing a test set and performance metrics), and outsource the entire data selection, collection and model building processes to the DEVA community. The community can then collaboratively build video analytics models, as follows:

- Everyone in the community can create, buy and sell data, annotations and video analytics models, using the DEVA platform tokens.
- Besides directly selling raw data, a data owner can also profit by purchasing an existing model, training/fine-tuning it with her data to obtain improved performance, and reselling the enhanced model using the platform tokens.
- A model builder, regardless of level of skill, can purchase a model, improve it with a novel design (possibly with purchased data), and resell it. Note that video analytics nowadays are mostly based on deep learning, which is largely a trial and error process. Performing such trial and error with the crowd is a cost-effective way of building an accurate model.
- The customer (in this case, the developer of a driver-assistant application) buys the final best models from the community.

Observe that in the above workflow data acquisition and model building are entirely decentralized and flexible – it is possible that no single party, including the customer, possesses the entire collection of datasets used to train the final models. Meanwhile, even with the whole dataset, it is possible that no single model builder can create the final models from scratch, since they are the collective work of the community.

Meanwhile, in DEVA the customer's job is very simple: define the problem and buy the best solutions. The entire process is integrated and streamlined. All value creation and value adding processes, such as data collection / annotation and model design / training, are entirely done by the community members in a flexible and not predefined way. The platform economic mechanisms through the use of tokens and other escrow services provide for the liquidity to perform the necessary transactions with minimum disruptions (need to access external escrow and payment systems).

Further, in DEVA, all data sharing / purchase occur between members in the community. No single entity, including the customer, is responsible for collecting and maintaining a dataset. In other words, the responsibility to ensure data confidentiality and privacy issues is distributed in the community.

Scenario 2: Smart Home

In this scenario, the customer aims to develop a smart-home video activity recognition app, such as one that recognizes patient falling, and alarms caretakers. The video data need to be captured in a home (for example, by a Nest Cam), and are thus highly private and sensitive. As a result, data owners are reluctant to share their videos, even with monetary incentives. The smart home maker would have to hire actors to perform these actions in a video-monitored room. After that, the smart home maker builds a model from the collected video samples. Can a developer for smart home analytics train an already available accurate model without collecting such sensitive videos?

Currently, only a relatively small amount of data is available in the public domain; the vast majority of existing data is either privately owned or cannot be shared because of privacy restrictions. This is especially true for data that contains personally identifiable information (PII), i.e., that allows the viewer or user to potentially identify a specific individual, such as video data for example. With privacy concerns increasing globally, and stricter privacy regulations being put in place by many governments around the world, our system makes it possible to train algorithms on datasets without running into issues with data privacy.

DEVA's solution to the above problem is based on federated learning, a new deep learning training technology recently invented by Google. Federated machine learning enables the training of models using data stored on other servers or devices. It decouples the model owner from the data owner and provides a platform for decentralized model training and testing on different datasets without ever having to share the model or data with anyone. Consequently, federated machine learning offers a distributed and scalable approach that enables better models and complete data privacy at the same time.

To further explain, let's consider a simple case with a model builder, who does not possess any data, and a data owner, who has data but does not want to reveal it to anyone. The model builder starts with a random model. Then, in each iteration, the model builder sends the current model to the data owner, and the latter updates the model based on her data. Note that here only models are circulated, and the owner's data is never shared with the model builders. This process can be expanded to involve multiple data owners, each providing an update to the model, (in the form of parameter updates) builder in every iteration. The model builder then averages these updates before applying them to the model.

DEVA facilitates federated learning through micropayments using the platform tokens: in particular, in each iteration of federated learning, every data owner earns a small amount of tokens. At the same time, throughout the process, no data owner reveals her private video to anyone else.

<u>Benefits for data owners:</u>

- Get rewarded and earn tokens for annotating your data and making it available for training.
- Data never leaves your own devices and is not shared with anyone.

<u>Benefits for model developers:</u>

- Access to a large number of data sources.
- Effectively train more performant models.
- Evaluate and test models on a larger and more diverse set of data.

<u>Benefits for model users:</u>

- Availability of better machine learning models.
- Assurance that models work on a large and diverse range of data sources.

Scenario 3: Medical Imaging Analytics

In the previous two scenarios, their goal is to build a single, end-to-end model trained from raw data. In some video analytics applications, it is common to involve a *pipeline* of models instead of a single one. Consider a scenario, akin to Kaggle's Second Data Science Bowl[1], in which a healthcare provider wants to build a model to estimate cardiac function (for example, in terms of end-systolic and end-diastolic volumes) from a video consisting of successive frames of MRI scans. Here, a typical solution involves two steps: first, to segment the heart chambers from individual frames, and second, to perform regression analysis on top of the segmented regions of interest.

DEVA facilitates video analytics pipelines by allowing model builders to freely trade individual sub-models of a solution pipeline. In the above scenario, an image segmentation expert could sell hart-chamber segmentation sub-models for DEVA platform tokens. A video regression model builder then buys such a segmentation sub-model, adds her own regression sub-model to form a complete pipeline, and re-sells the final solution. In general, there can be different possible pipelines, for example, one with 2-chamber segmentation and another with 4-chamber segmentation. DEVA as a platform does not dictate the pipeline; instead, the market, driven by demand and supply, directs efforts spent on building and refining different sub-models as well as pipelines.

Benefits for data owners:

- Get rewarded and earn tokens for annotating your data and making it available for training.
- Data never leaves your own devices and is not shared with anyone.

Benefits for model developers:

- Access to a large number of data sources.
- Effectively train and trade sub-models.

Benefits for model users:

- Availability of better machine learning models and video analysis pipelines.

[1]https://www.kaggle.com/c/second-annual-data-science-bowl

Implementation

We break down the necessary developments into the following components, details of which are provided below.

- Video analytics algorithms/models
- Economic architecture
- Security and privacy
- Auditability/proof of integrity of contracted workflows
- Testbed

Video Analytics

The proposed platform will enable a structured approach to video analytics. Its purpose is to enable modular and configurable video analytics pipelines to support all stages of video analysis while promoting the re-use of various components and intermediate results. These include the following:

- Video sources: The platform needs to support processing and combining video from many different heterogeneous sources: for example, cameras of different capabilities/quality, mobile, in-car and wearable cameras.
- Video annotation: Labeled ground truth data is essential for most data-driven and learning-based approaches for video analytics. The platform will support user-based video annotation and labeling.
- Deep learning models: The platform will support training, tuning, and testing of deep learning models and architectures.

We will seed the platform with simple versions of all of the above (see also the Section on Testbed).

Economic Architecture

We design the necessary economic components that are required for the operation of the platform. These include flexible smart contracts, payment functions and escrow services via DEVA platform tokens, a service market place with publish-subscribe services for suppliers and consumers, brokers for assisting customers to find/build the appropriate service bundles. The goal is to create an environment where multiple agents can cooperate in auditable and flexible ways to trade services, or jointly build and sell services to other agents. The market place contributes in discovering and pricing services, generating the appropriate revenue streams while sharing fairly infrastructure costs, and ensuring the sustainability of the platform.

A fundamental challenge is to design a portal serving as a market place for video services and metadata. Suppliers will provide economic factors like raw video data or access capabilities to cameras, modules for analytics, metadata and post-processed information. Following the scenarios mentioned earlier, consumers will post requests for video metadata specified in our ontology framework. Specialized brokers will decompose these high-level requests into requests for more basic services that suppliers in the ecosystem can provide and will either elicit the corresponding bids or use prices from a pre-specified catalogue. Alternatively, participants may cooperate in the creation of competitive models using jointly acquired resources from other participants. Flexible smart contracts will define the sharing of costs and revenues from customers that access the new services through the market place. The market place will include reputation services and transaction histories that can reduce information asymmetry and facilitate interaction between participants.

Security and Privacy

Security and privacy are important factors to the success of our proposed infrastructure. Curious users or malicious attackers may attempt to exploit it to retrieve or infer personal sensitive information. In traditional IT systems, the accessibility to personal information is usually directly measurable, which allows systems to enforce security and privacy protection by data-oriented access control and query output risk evaluation. In an open video analytics marketplace, however, sensitive information regarding individuals, e.g. personal traces, may be present in the visual data. This raises challenges to maintain privacy protection of the data at all times.

DEVA's solution to confidentiality and privacy protection to make extensive use of Trusted Execution Environment (TEE) technologies, such as Intel's Software Guard Extensions (SGX). In a nutshell, TEE allows trusted code (which can be open-sourced, audited, and signed) to be executed in an untrusted computing platform, which includes the operating system, the hypervisor, the BIOS, and all hardware except for the TEE itself. The trusted code is executed in a *secure enclave*, which is attested by the TEE manufacturer, e.g., Intel's Software Guard Extensions (SGX) or AMD's Secure Memory Encryption (SME)/ Secure Encrypted Virtualization (SEV). Each secure enclave directly communicates with other secure enclaves (possibly over the Internet) through encrypted channels. TEE ensures both confidentiality of data (i.e., the inputs) and integrity of computation results (the outputs).

In the context of video analytics, TEE enables a blockchain verifier (e.g., a miner in a Proof-of-Work chain, or a validator in a Proof-of-Stake chain) to evaluate the performance of a given model on a given test dataset, without revealing either the model or the data to the verifier. This is done by executing audited code in a secure enclave that applies the model on the data. Meanwhile, TEE also guarantees that the evaluation result, i.e., model performance, cannot be falsified by a dishonest verifier. Such evaluation results establish the value of the model, and are a key input to the smart contract transactions that trade models for DEVA tokens, e.g., a buyer can place a bid on a model with the condition that it achieves 97% prediction accuracy, and processes 30 video frames per second. The availability of evaluation results also enables comparisons between different models, and avoids a lemon market in which buyers are unwilling to place high bids due to uncertainty about the value of the model.

Similarly, TEE also enables a block verifier to evaluate properties of a dataset. For example, in some application scenarios positive samples (such as videos of car accidents) can be more valuable than negative ones, and diversity of the data attributes (such as videos captured in stormy weather) can also be an important factor. Such properties can be verified by trusted code in a secure enclave, without revealing contents of the dataset to the block verifier.

Testbed

As part of the DEVA platform efforts, we are developing a flexible testbed infrastructure that can support video data from different stakeholders. We will contribute various proprietary and open-source video analytics models and datasets to seed the platform.

Timeline & Roadmap

The milestones and roadmap for the development of the DEVA platform are as follows:

- Milestone 1 (T+6 months): basic platform and features (scenario 1)
- Milestone 2 (T+9 months): federated training (scenario 2)
- Milestone 3 (T+12 months): host competitions (populate platform)
- Milestone 4 (T+18 months): vertical compositions (scenario 3)
- Milestone 5 (T+24 months): supporting data beyond video

Time T indicates the end of the token sale (see below).

Token Sale

DEVA tokens minted: 1,000,000,000

Conversion rate: XXX DEVA tokens created for 1 ETH

Timing:

- Private sale starts 1st January 2019
- Public sale starts 1st April 2019
- Sale ends 30th April 2019

Token distribution:

- 60% issued in token sale
- 15% reserve
- 15% to the team
- 10% to partners

Proceeds distribution:

- 50% development
- 20% services (cloud, legal, etc.)
- 10% marketing
- 10% office and equipment
- 10% to founders and backers

Team

The team behind DEVA consists of renowned scientists and experienced entrepreneurs.

Andrew Ow has 20+ years of experience in roles covering business development, marketing, operations and management in the Infocomm technology industry. He started an SBO business in 1999 and sold the company 5 years later. His background includes regional partnerships with Telcos, system integrators, distributors, OEMs and independent software vendors. He also has chaired many volunteer committees in Singapore.

Dr. Stefan Winkler is an established research leader and entrepreneur with 20+ years of experience in industry and academia. He co-founded two companies with a focus on video analytics (Genista and Opsis). He is an IEEE Fellow and holds a Ph.D. degree from the Ecole Polytechnique Federale de Lausanne, Switzerland, as well as a M.Eng./B.Eng. degree from the Technical University of Vienna, Austria. His expertise includes image & video processing, computer vision, machine learning, human-computer interaction, user experience, and perception modeling. He has published over 120 scientific papers.

Dr. David (Yin) Yang is Assistant Professor Hamad Bin Khalifa University, Qatar. He has a Bachelor of Engineering from Shanghai Jiao Tong University and a PhD degree from the Hong Kong University of Science and Technology. Dr. Yang is an expert in big data analytics, data security and privacy.

Dr. Tom (Zhengjia) Fu is Research Scientist at the University of Illinois' Advanced Digital Sciences Center, Singapore. He has a Bachelor of Engineering from Shanghai Jiao Tong University and MPhil and PhD degrees from the Chinese University of Hong Kong. Dr. Fu is an expert in distributed stream data analytics, cloud computing, and peer-to-Peer content distribution systems.

Dr. Costas A. Courcoubetis is Professor at the Singapore University of Technology and Design (SUTD) where he heads the Initiative for the Sharing Economy and co-directs the new ST-SUTD Center for Smart Systems. He holds a Diploma from the National Technical University of Athens, Greece, as well as MS and PhD degrees from the University of California, Berkeley. His current research interests are economics and performance analysis of networks and internet technologies, sharing economy, regulation policy, resource sharing and auctions. He has published more than 100 scientific papers, with over 13,000 citations. He is co-author of "Pricing Communication Networks: Economics, Technology and Modeling" (Wiley, 2003).

Dr. Georgios Piliouras is Assistant Professor at the Singapore University of Technology and Design (SUTD). He holds a PhD degree in Computer Science from Cornell University and is the recipient of a Singapore NRF Fellowship. His main research interests lie in the areas of algorithmic game theory, computational learning theory, multi-agent learning and dynamical systems. His current research focuses on the intersection of game theory, blockchain, and cryptoeconomics.