OPEN AR CLOUDs’

STATE OF THE AR CLOUD REPORT

OPEN AR CLOUD

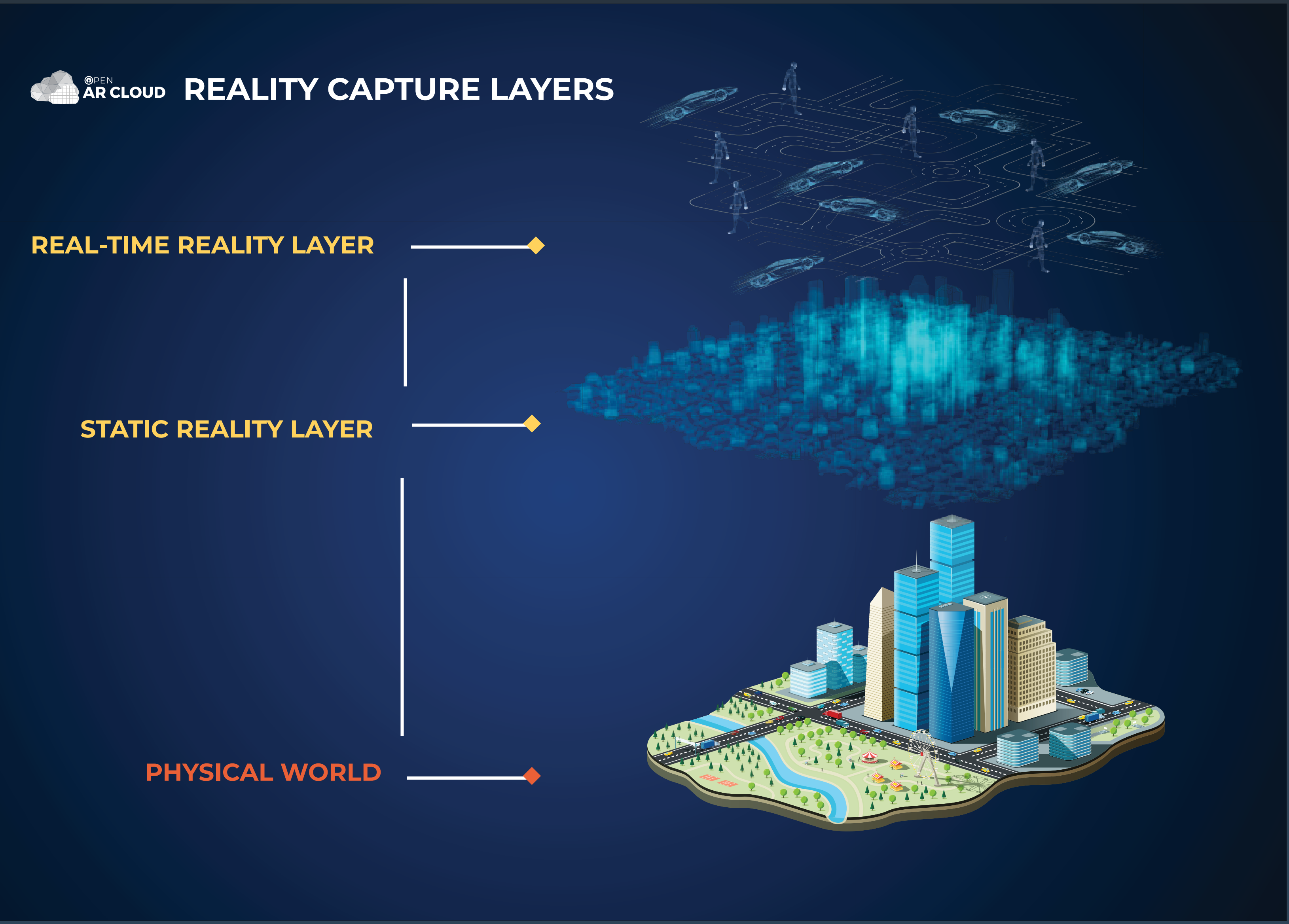

OPEN AR CLOUD
REALITY CAPTURE LAYERS
REAL-TIME REALITY LAYER
STATIC REALITY LAYER
PHYSICAL WORLD

Table of Contents

About This Report

This document is a little piece of history. It's the first production of the newly founded global non-profit Open AR Cloud association (OARC). It has been written collaboratively by leading subject experts from around the globe who are taking part in OARC's many Working Groups.

It is also the first report that provides an initial consensus overview of the state of many components that make up the nascent AR Cloud industry that will be at the center of the next big platform shift in computing, bringing the internet off screens and out into the real world.

The report lays no claim to being comprehensive and cover everything, but should provide an unprecedented overview of what is happening, and cover a wide range of topics regarding the birth of real world spatial computing. It covers the core components that will enable this new digital reality. It details the state of the industry, the companies involved, the wide range of technologies and solutions in development, and the myriad of challenges to overcome, with special emphasis on the interoperability, scalability, usability, safety, privacy, ownership, security, compliance and legal issues ahead.

It also attempts to analyze those challenges regarding building an AR Cloud ecosystem (also referred to as the "spatial web"). We intend it to educate and inspire individuals, technology executives and policy makers to consider the challenges ahead carefully and address them based on specific values and principles that will best benefit all of mankind.

Executive Summary

The current state of the AR Cloud is: promising, early and incomplete.

We are at the beginning of the era of real world spatial computing, an evolution of our digital lives we expect to grow to be at least as central to humanity as the web and smartphones are today. It will focus around a set of technologies that introduces seemingly insurmountable technological, privacy and societal issues. How and when we will address those issues is still unclear.

What is clear to the authors is that this is the dawn of a computing paradigm where the digital world will fuse together with physical reality. This real world "spatial web" will transform every aspect of our economy, culture and society, and impact every aspect of how we live, work, learn and play. "AR Cloud" technologies will be a fundamental prerequisite for this to happen and provide the required scaffolding for the shared programmable space of this era.

This is the year where millions of people will experience the first glimpses of AR Cloud technology for the first time. The most prominent example at the time of writing is AR navigation in Google maps which works because their “Visual Positioning Service”, an example of AR Cloud technology, provides precise and accurate position and orientation of the device on a level that GPS and compass could never deliver. We expect several more examples to appear this year. We have provided a long list of companies currently working on bringing AR Cloud technology to the market in this report illustrating some momentum behind what is about to happen.

Our goal is to help the industry grow and work together through a common understanding of the terms and issues ahead. We aim to develop a flexible conceptual framework, encompassing technologies using clear terms, standards, policies and best practices that will guide the development of each of the core components required, to create an open and interoperable AR Cloud ecosystem.

Also included is OARC’s vision for an open and interoperable real world spatial computing platform, that will best serve all mankind, at the dawn of this new era. Central to this vision is the belief that AR Cloud based content, experiences, and solutions must work seamlessly on any platform or device, for all people, anywhere in the world, ensuring the protection of user privacy, security, safety and welfare above all.

Key Findings

Creating a real time representation of the world filled with shared digital content and experiences is vast and thus each chapter of this report contains valuable detail readers should not only review, but get involved and improve.

1. **This is significant**
   Spatial computing will fundamentally change every aspect of how you live, work and play. Tim Cook, the CEO of Apple called its impact, “profound” in 2018.
2. **This is happening. now.**
   Companies such as Google, Apple and others are spending billions to adapt and build required spatial computing infrastructure taking many forms. Entrepreneurs are building initial systems and early examples of both

commercial and consumer platforms and applications that are paying dividends today.

3. **Reality capture will take place in real-time at unprecedented detail**
   Continuous, consistent reality capture is a big goal and has already begun in private and public spaces. Many technology providers are making big bets on this.
4. **Big data and machine learning is at the core**
   The big data and artificial intelligence technologies, namely machine learning, contribute in many fundamental ways to the indexing of 3D maps for AR Cloud services.
5. **The technological challenge is vast**
   Creating a multi-dimensional distributed real time representation of the world is an enormous task and will require technologies, architectures, devices and methods we have yet to develop.
6. **The personal and societal issues are greater**
   Without trust and respect for users, their data, intellectual property and intentions, we will never reach the full potential, adoption or acceptance by society. And without a deep understanding of compliance, legal, regulatory, and even national security issues companies could quickly lose their license to do business.
7. **We're missing key components, standards and infrastructure**
   Comprehensive, global, real time, machine readable maps required do not exist. The ability to sense our dynamic world and put it in context is raw, but promising. We must develop guidelines for secure and privacy-protecting ways to locate users and devices and we should standardize the output. We must develop our ability to create, share and publish digital content and experiences...the list goes on.
8. **Content and delivery will be crucial for consumers**
   We will need to both reach a critical mass of content and experiences, and it needs to be accessible from spatially and thematically indexed registries (that does not yet exist), for users to embrace this new computing paradigm.
9. **Low latency high throughput edge computing will be key**
   End user devices cannot keep up with the processing requirements and the big data, but we cannot realize many important use-cases if latency is too high. Hence, delivering AR Cloud-based services and experiences will heavily on the development and standardization of new edge computing architectures
10. **The opportunity for mankind is without equal.**
    We cannot understate how this new digital reality will unlock new levels and qualities of human existence, relationships, knowledge, entertainment and productivity.

11. **Its future implications are profound**
    A natural interface to the real world and our imaginations will propel mankind to new heights. AR Cloud systems have the power to change the definition of "reality" for a new generation of imagi-natives, born into a world where they can speak their mind into view, where bits and bytes are as "real" and viable materials as wood, steel or stone.
12. **You are not too late**
    If you are reading this, you are witnessing the birth of a new era of unbounded opportunity. It's times like these when the next Google or Facebook can emerge, so dig in, read on, and imagine a better future in the world around you…then go create it!

Introduction

A highly motivated and capable community of technologists works feverishly to change the way humanity sees, contributes to, and uses information forever: to unite the digital and physical worlds. The impacts of implementing Augmented Reality well, fairly and globally are so big that they have been the focus of writings, movies and other culturally relevant works for decades and maybe longer.

Where there is potential for great impact, there is also a high potential for mass scale misinformation, deception, confusion and misunderstandings. This report is the first collaborative survey and analysis of AR Cloud and many underlying, enabling technologies which both captures facts and provides thought leadership. This report captures the state-of-the-art of AR Cloud at the time of its publication, and to contribute to the advancement of a shared vision.

This report is unique in perspective and scope. The perspective is solely that of the members of a young association whose mission is to advocate for the Open AR Cloud ecosystem built upon key principles. We introduce the association's origins and history below, and describe the organization's founding principles and vision in further detail in the following section. The report's scope includes everything that could impact the smooth and rapid success of the AR Cloud: from technology to policy, from protocols to human factors.

The contributors to this report learned a great deal while conducting research and compiling their findings into one complete resource. The preliminary nature of some sections of the report show that the work must continue; the OARC expects to release regular updates to this report. But, the full and future value of this report will ultimately be measured by its readers who will gain and soon have the opportunity to use new insights. They will learn about the many enabling technologies upon which the AR Cloud industry will build and the challenges ahead which must be solved if the AR Cloud is to meet or exceed its potential to benefit humanity.

The Organization and People Behind This Report

The Open AR Cloud (OARC) is a global non-profit organization founded in 2018 with the mission to drive the development of open and interoperable spatial computing technology, data and standards to connect the real and the digital worlds for the benefit of all.

The seeds were planted in the spring and early summer of 2018, by fostering discussions at a panel during Augmented World Expo in Santa Clara. The panelists, and dozens of others who have formed and driven multiple earlier initiatives and pioneered companies at the forefront of the field, shared a sense of urgency. They saw a great potential for good, with many risks. Over the following months, they joined forces to work together for a better future.

Their shared conviction is that a unified organization will be better able to bring together forces to drive the future direction of AR Cloud technology. The organization encourages approaches, standards and best practices that will bring the greatest benefits to the most people while also avoiding potential threats to privacy, security and safety.

Launched on October 28th, 2018, at the Augmented World Expo EU in Munich, where it released a draft *Privacy Manifesto for AR Cloud Solutions*[1], the OARC brought together over 240 founding members and 36 founding partners within a few weeks. Since then the number of members and partners has steadily continued to grow. This report is the first product of those passionate professionals from around the world who have contributed towards the first of many objectives: to educate all audiences about the Open AR Cloud.

On the 21st of February 2019 the organization, under the leadership of its Governing Board, held its Inaugural Working Group meetings. Dozens of members and partners from around the world took part in developing the first charters of and selected leadership for each of the working groups. This represented a historical milestone in bringing together pioneers in the industry for the purpose of building an Open AR Cloud ecosystem as a collaborative endeavor.

[1] https://medium.com/openarcloud/privacy-manifesto-for-ar-cloud-solutions-9507543f50b6

One responsibility assigned to the working group members, as part of their forming cohesive groups with clear goals, was to generate a clear and complete overview of their understanding of the AR Cloud from the point of view on which they were to focus. Shortly after the kick-off, the leadership agreed that the OARC would aim for its first deliverable to the community: one complete report compiling each working groups' assessments and recommendations for future work. This report is a snapshot of the current best knowledge and insights the working groups have gathered. It is being distributed openly and freely to the world and shares what the OARC members have learned.

The Invitation

The Open AR Cloud results from pioneers who have led prior initiatives within the nascent AR Cloud industry sharing a sense of urgency and deciding to join forces to drive the development of this technology for the benefit of all. The founding members and partners want to ignite a digital renaissance in which real world spatial computing technology serves every person. They also share the conviction that all the stakeholders are part of the solution and should be included in working towards the vision.

Collaboration is essential and highly encouraged. It can be at any level: leadership, grassroots, communications, coding. We invite individuals to become members. We encourage companies to become partners. All are valuable and important to achieving the desired outcomes.

Unique Value

The preparation of this report has been a collaborative effort to capture the current status of many disparate technologies and trends. In this respect, it is entirely vendor neutral and presents, to the extent by those who collaborated on it, a consensus view of AR Cloud in mid-2019.

The OARC is an open, inclusive organization. As its first public product, this report seeks to be entirely vendor neutral. We based all our information about technology strategies or solutions on public sources and provided it without proprietary agenda. This said, technologies and tools are rapidly changing. We provide all information without warranty and the reader, upon use of the contents of this report, may not hold the OARC liable for any outcomes.

While vendor neutrality is a key value, thematic focus and perspectives taken by the authors were inspired by the vision OARC has for the future: for the digital and physical worlds to connect for the benefit all. The proposals and recommendations expressed in this report are, at their core, offered with the goal of promoting interoperability and the formation of a vibrant and active ecosystem of companies, each providing value to customers. They were also designed to discourage the formation of closed technology silos, or the emergence of forces that inhibit competition, expose personal data, and harm businesses or users in other ways.

One of the important steps towards achievement of OARC's mission is to develop and to use a common vocabulary. At the end of this report there is the first draft of a complete glossary of AR Cloud terminology. Many of the terms included involve components of the AR technology stack and other enablers.

Many concepts underlying AR Cloud and related topics are new. Language within the industry, let alone amongst the general population, is often unclear. Some concepts discussed in this report and in the industry do not have agreed upon terms or definitions for commonly used (or misused) terms. By using definitions proposed in this glossary, and offering/defining new definitions for these or new terms, stakeholders will reduce confusion and accelerate widespread understanding of AR Cloud topics.

The glossary exists as an open repository on Github to allow third parties to contribute to and discuss the terms and definitions in the glossary[2].

How to use this report

The primary purpose of this report is to decrease confusion among stakeholders and raise awareness about the important unsolved challenges and critical topics.

In order for conversations about AR Cloud to be productive beyond the small existing community, the OARC seeks wide distribution for this report. The more we discuss the concepts openly, the better. Whether to agree with and support the mission or to correct its shortcomings, we encourage readers to share any part in print or online.

The following citation must accompany any quotes of this report or usage in any media: "From the first State of the AR Cloud Report published May 28 2019 by the Open AR Cloud. All rights reserved."
When quoting or citing this report on the Internet, include a link to the Web site for the report: http://stateofthearcloud.com/

[2] https://github.com/OpenArCloud/OpenGlossaryForRealWorldSpatialComputing

To stimulate dialog with large and diverse audiences, we highly encourage readers to use social media, to join the OARC Discord server and to contribute to online forums where OARC members and partners can provide feedback and recommendations.

As time and resources permit, the website for this report will host additional supporting materials but at this time, all are encouraged to join OARC.

AR-enabled devices and AR Cloud technologies seem poised to at the center of this convergence, giving humans the ability to experience together and interact directly with the digital world and virtual objects in our physical surroundings, using all our senses and the full range of capabilities of the human body. There is a tangible sense of excitement around how this technology could soon usher in a new era where all aspects of how we interact with the digital world, each other, and our physical surroundings will radically improve. AR Cloud technologies could be the key to unleashing untold new capabilities and opportunities for all of humanity to prosper and lead more interesting and fulfilling lives.

But by the nature of how it works, these technologies can, if society is careless, threaten our privacy, security and even freedom in unprecedented ways. Instead of unleashing human potential, we could turn it into a digital prison from where escape might be impossible.

Nobody can predict the details of such a complex and constantly shifting and unformed technological landscape far into the future. Although the OARC community brings to bear a lot of varied expertise and experience, and even though we wish to boldly explore and help shape the way of the future, we realize there are many known unknowns and even more unknown unknowns. We wouldn't at this point in time know what sort of approaches and solutions will be optimal 5 or 10 years from now.

But there are some solid concepts that we must adopt, and fundamental values that could guide those navigating the unknown.

Before examining the current status of AR Cloud, the following section will explain in detail the OARC vision for the future development of the AR Cloud ecosystem.

Our Vision - An Open and Interoperable Ecosystem

As made clear by the name of the OARC, the concept of openness is one of our core values. We believe the open approach is what will ultimately give the best overall long-term results for the vast majority of stakeholders, whether they are users, developers, governments or anything in between.

Content, experiences, services and solutions should seamlessly deliver value to all users, regardless of their choices of access device, AR platform, or geographical region.

Proprietary technology silos, or "walled garden" ecosystems, have benefits in some circumstances, however, they also introduce problems, the most notable of which are vendor-lock in, slow or low innovation, limited market reach and reduced transparency. How the open web platform has and continues to permit content creators, service providers and developers to create their experiences and solutions once and then to reach an audience of 4 billion people across the globe who can experience their creations on any browser on any device is inspirational to the OARC. Open and interoperable systems for information delivery are also highly beneficial to the users. Today, users with the web browser of their choice can safely access a much greater variety of content and solutions than they would be if we locked them into one or a few walled gardens.

For open and interoperable AR Cloud solutions and services to be available to users, the ecosystem needs many, collaboratively developed and widely implemented standards, protocols and evolved shared, globally scalable infrastructures. One concept that might be transferable from the Web platform is a t browser for the digitally enhanced physical world.

If we are to “paint the world with data”, it would be of great benefit if there was a standardized way of fetching and presenting all the different colors of paint to the users.

A decade ago, when early mobile AR was based on printed markers or simple GPS location and compass, there were visionary efforts to create a standard for AR browsers. Sadly, those efforts were stranded. But the idea has merit and the potential payoff of delivering on the promise within the current technological landscape presents even a greater opportunity.

OARC envisions a form of “Universal Spatial Browser for the Real World” and we think it might be even more important to the next era of computing than the web browser is for the current paradigm. This report presents an early list of prototype browser with some rudimentary capabilities. OARC is following with great interest the efforts of groups such as the W3C’s Immersive Web WG and hopes to engage more with that effort going forward. For the best real world browser, it is unclear if it would be better to start from scratch or to evolve the current browsers to gain new capabilities. Just to figure out the initial requirements for a fully featured browser for the real world requires a large collaborative effort. Another body of work could be to develop standards that enable spatial content, experiences, services and communication to work seamlessly and safely on any device and that any presentation software or “viewer” could understand consistently. The OARC members and partners are contributing towards the goal of advancing open and interoperable systems through their support of many working groups.

Openness and Inclusivity in How We Create The Ecosystem

OARC welcomes participants from around the globe to help create this ecosystem in a way that is transparent and where the organization will share what its members and partners contributes with the world. OARC also want to involve a wide range of stakeholders in different areas.

The process should be open, transparent and involve representatives of all demographics, from all parts of the world, from all walks of life, with people both inside and outside of tech.

The original web was the brainchild of Sir Tim Berners-Lee who single-handedly drafted the concept of the web in a one page diagram 30 years ago. The primary use case he envisioned for the technology was the more efficient sharing of scientific textual information that could refer to other scientific texts using hyperlinks.

Over the years as the web platform evolved to be an integral part of most aspects of modern life, developing and expanding the platform gradually started to involve more people in an increasingly open way. However, technologists control this process with rather limited involvement of outside stakeholders. Thus it also lacks the diversity of views typical of other technology domains.

The architects of the AR Cloud can and should aim for better. The era of real world spatial computing is for everyone and leaders must involve people in all walks of life, both inside and outside of technology and including all demographics and geographies. Everyone should have a say and be able to affect the future of a technology likely to affect in profound ways how they live their lives.

Ultimately, the technical implementation details will be up to “the techies” and other experts, but this time around they should work tightly with people representative of the rest of humanity to evolve the requirements for the specs and protocols.

OARC as an organization intends to establish the widest stakeholder involvement in its working groups and other activities, and to have transparency in decision making. The beginning is promising, but there’s a long way to go before the diversity of representation is at the level it should be.

Promoting Development of Best Practice UX Design

OARC believes that we should design the user interfaces for AR Cloud technology to serve all humans, especially those with differing needs and abilities under varying circumstances. Content, services and experiences should be designed from the bottom up to be accessible, user-friendly and safe for everyone.

We expect AR Cloud technology to increasingly harness and use machine readable representations of the real world as well as to offer content and services in a spatial context. OARC envisions that such capabilities could enable a radical improvement in inclusive human machine interaction and UX design. We even believe this technology could make the real world much more accessible to people with disabilities like lack of eyesight, hearing or mobility.

That AR technologies augments visual, auditory and other sensory inputs to the user as the user is moving through the real world, introduces a range of user safety issues that could affect anyone, regardless of individual needs and abilities. We must take special care to ensure systems do not cause users to get distracted or disoriented in risky situations like near a road with heavy traffic, next to a cliff or a raging river.

Establishing best practices for entirely new types of experiences and solutions requires systematic and resource intensive work that includes a diverse range of people with varying needs and abilities through the entire design process. OARC already has a robust UX community on a path to explore the first steps towards developing such guidelines. There is definitely a need for funding and we want to collaborate with the industry and academia on advancing this new field.

Once we have developed initial best practices and guidelines, OARC will do our best to promote the application of such practices and guidelines in the industry, while also promoting a continuous process of testing and iterating over the best practices as the technology develops.

Standardizing Outputs of Reality Capture and Discoverability

We predicate the core functioning of everything AR—from the most basic of AR Cloud technologies to much more sophisticated, future capabilities—on the ability to capture reality and transform the captured data into something that makes the geometry, semantics and context of the world machine-readable. This is an evolution of a long history of mapmaking and will represent a big advancement in mapping technology in the level of detail available, the rate of which we can update maps and embed the richness of context in them. Every time there has been a significant improvement in mapmaking technology humanity has experienced economic growth and increasing sophistication of civilization.

Although, as is clear from this report, the methods and sensors of capturing data about the real world can differ significantly, they all should inform and update convergent shared representations of the world, standardized from the point of view of the end user. It is OARC's opinion that one should build upon and expand existing standardized schemas of digitally representing the world as a base case. That is why reviews of standards from organizations such as Khronos Group, W3C, OGC and buildingSMART are already underway. If we need more radical improvements,OARC members or partners will seek to work together with the most appropriate Standard Development Organizations (SDOs) to propose new standards or extensions and update to existing specifications.

A feature of a machine-readable world that OARC wants to bring attention to is that it has two fairly distinct parts or “reality layers”:

1) **The mostly static “base-layer” of 3D terrain, buildings, roads, trees and even interior walls**
2) **The “real-time layer” containing context rich representation of people, vehicles, animals and movable objects**

Although, from the perspective of the user, these two parts will together create a whole, OARC envisions that the capture and handling of those two reality layers will be rather different.

Although the static layer may contain information that should not be available publicly or shared such as details of a military facility, the layout of critical parts of a nuclear power plant or the interior walls of private homes, we could regard most of the data in this static layer as “public” and share it across the globe.

For the temporally dynamic reality, things are very different. Making this data machine readable through technologies like “semantic scene understanding” will be very useful for anything from helping machines and people better respond to complex dynamic traffic, playing and working together using AR-cloud technology, help to aid people in unfamiliar situations and making the world more accessible to people with a range of disabilities.

On the other side, this data is of an intrusive nature, and even current machine learning capabilities would, if given access to the data, be able to read it to understand and predict with unprecedented accuracy the likely behaviour, a state of mind, health, interests, attitudes of individuals and their relationships with other people.

It is OARC’s position that no organization, commercial or governmental, should be allowed to collect and aggregate this data centrally for large areas or the entire world. Such aggregation would represent to big a risk of abuse of power over the citizens. We base our position on the fundamental values of respecting privacy, dignity and freedom of individuals. These values are also expressed in the OARC draft *Privacy Manifesto for AR Cloud Solutions* that was published on the 18th of October, 2018.

There is benefit in sharing data in the real-time layer locally and OARC proposes that we should enable such sharing in controlled and highly secure ways only on the edge and by default go no further. If we do not create and respect this real time reality “firewall”, citizens and users may lose all sense of privacy, freedom and dignity.

Communication between people at different locations should allow some context to transmit in highly encrypted forms and, importantly, the parts of the context containing personally identifiable data should, by design (unless explicit and scope-limited permissions are granted by the individuals), only be accessible to the people who take part in the communication. And the receiver of such data should not, by design, store such data without giving notice. There are today automated ways to facilitate filtering of such data and we could adopt these for any AR Cloud-based communication solutions.

Also, the safety and security of citizens would be at risk if bad actors can monitor their activities and context in such great detail. This technology introduces a lot of potential threats that we must manage carefully.

There are situations that might give rise to exceptions from this edge-only principle, such as natural disasters and terrorist attacks. However, it is OARC’s opinion that transmitting and storing any such data should be highly regulated and subject to third-party control. When there are situations where government agencies need to breach into this data there should be transparent accountability without unreasonable delay on each event so that citizens can monitor the level of surveillance their authorities have been enacting.

Respecting Local Regulatory Policies and Personal Privacy Choices

OARC upholds the human rights and democratic principles and is of the opinion that in every community and nation it is “We the people” of those places who should ultimately decide the law of the land. The OARC will seek to enable its envisioned Open AR Cloud ecosystem to be flexible and adaptable to the will of the people, empowering people at a local level so they can implement the laws and regulations they see fit.

For those areas of the world where democratic rule does not exist, technology providers and hosts must refer to principles of universal human rights to inform how these organizations use the technologies. But OARC has no jurisdiction over countries and does not seek to obtain the power and ability to enforce such principles. Just like we use web technology in both democratic and undemocratic regions of the world and where some populations are subject to different levels of censorship and government surveillance, OARC does not expect that providers and users will be able to avoid similar scenarios for AR Cloud technologies entirely.

We should note that democratically governed regions should take measures to protect themselves and their citizens from intrusions coming from regimes who would seek to compromise their stability. Use of technologies always pose threats, but AR Cloud technologies could provide new ways for surveillance, hacking and manipulation.

Regarding individuals' right to privacy, OARC sees legislation like the GDPR as a move in the right direction. It believes that individuals should be in full control over all their personal data on as an inalienable right.

Just like slavery was abolished and it was made illegal to own and sell other human beings, no one should sign away their right to the future control of their own private data and that permanently owning other people's personal data should be made illegal. It should always be up to the individual to decide with whom to share what types of personal data, to be provided with means to know about how the data is used, to have the ability to prevent its uses and pull the data back from the other party.

Empowering Flourishing Local Communities and Economies

Most of the data related to AR Cloud and real world spatial computing is most valuable at its origin. Technically, it is more optimal to store spatial data at the edge, as it reduces latency and data transfer costs. In fact, many important future use cases such as a real-time shared representation of moving vehicles and pedestrians, can only work and save lives if we handle the data locally.

But this decentralized approach offers a great opportunity for local communities and businesses and individuals to regulate, create and control local data, content and services.

This could help spread self determination and empowerment more widely in the world and could mark a healthy shift away from the current extreme concentration of power, data and wealth amongst the largest tech companies.

Local players that create and provide services and content to local customers might have little need of 3rd party organization to mediate the discoveries and transactions thus keeping a higher share of the economy within the local communities. An added benefit of buyers and sellers being able to find each other locally on their own terms is that it might provide less of an incentive for third parties to spy on and predict the behaviour of individuals.

OARC hopes that within a few years there will be many examples of how the Open AR Cloud ecosystem we envision will allow local businesses, or communities to provide local content and services. By leveraging open and standardized protocols, common data formats and spatial and thematic indexing schemes, there will emerge new services for automated AR experience discovery by customers. Annotating content such as GLTf files, subscribable services for the accurate position and orientation of local public transportation vehicles and other things with geopose and listing it in standards compliant local indexes, might be one first step in this direction. Universal standards for facilitating transactions for purchasable spatial content and services would be of enormous value as well and incentivize content creators.

We also believe that Distributed Ledger Technologies and decentralized apps (dApps) could play an important role in facilitating such capabilities. That a reason why we have set up a Working Group to explore those possibilities.

Accelerated Technological Innovation Through Open Data, Open R&D, and Open Source Code

OARC's commitment to Open goes beyond the collaboration on the standards, protocols and shared infrastructure. AR Cloud technology is inherently complex and from what we see the capabilities that the industry provides or plans to bring to the market soon are rudimentary. There is a long way ahead of us before our industry can reach some of the capabilities we envision for the future.

OARC is in no position to advise startups to give away their competitive advantage for free, but we recognize how open source, open data and open R&D is behind much of the prosperity and financial successes of our time. We can attribute part of the 2012 Deep Learning "big bang" cto how researchers in the field publish their breakthroughs on open access sites like arXiv.org, their source code for frameworks and neural net models on sites like GitHub, and the training data in shared repositories like image-net. Another example of how an open approach helps an industry flourish is how the release of a lot of Tesla Motors patents for Electric Vehicles has helped the rapid growth of the EV industry.

The OARC point of view is that the AR Cloud industry could benefit from projects such as collaborating on an open source R&D toolchain for simulating the performance of AR Cloud technologies based on computer game engines. A state-of-the-art simulation of the real world, plus simulation of the device sensor streams, the emulation of the device processing power, and data transfer over networks and edge or cloud services would enable developers and research institutions to test out their algorithms under varying conditions quickly. We could use parts of the same toolchain to create much needed pixel perfect training data made publicly available.

The autonomous vehicle industry has proprietary versions of those types of R&D toolchains, such as NVIDIA's DRIVE Sim, and internal ones at Waymo or Tesla Motors. The main challenges right now is that no such toolchain exists on the market or as open source for AR Cloud technology and that one cannot simply copy+paste the ones for autonomous vehicles. AR devices are far more constrained in processing power and energy use, and it has a different set of sensors than self-driving cars. You would need to have an R&D toolchain that lets you tune the emulation of the amount of processing that can happen on evolving generations of AR capable devices versus the processing that could happen on the infrastructure edge or in the public cloud. OARC believes developers can create such a toolchain based on open source game engines like Godot[3] and is seeking to establish such a project together with industry partners and sponsors.

Another open source project that could be widely useful would be a reference implementation of different base functionalities like conversion libraries between a 6degrees of freedom geographical position and orientation and the Cartesian x/y/z coordinate systems used for AR SDKs like ARCore and ARKit.

It is in this creative process of innovation where open source, open data and open R&D could play a vital role in both OARC's work, and for the next era of computing.

[3] https://godotengine.org/

This chapter analyzes the AR Cloud enabling technologies and principles introduced in Vision of Open AR Cloud and Framework for Open AR Cloud, above. Thought leaders and subject experts volunteered in OARC Working Groups (WG) to contribute the following sections.

Reality Modeling and Mapping

Introduction

Reality modeling and mapping is the process of: capturing the spatial map of a real space, determining the geometric relationships between objects in that space, modeling the state of those objects and the semantic relationships between them, gathering it all into a digital collection.Then, storing that digital collection against an index, so that it is easily queryable.

This section covers many of the technical topics relevant to the creation of spatial maps and semantic scene understanding in the context of the AR Cloud, such as computer vision techniques, current vision-based mapping and location technologies and various sensors that can be used.

Scope

Reality modeling and mapping is one of the most critical components of the AR Cloud ecosystem. The output of this process underpins the entire set of systems that comprise the AR Cloud by providing data that describes the structure and the state of the real world. This data is used to position the user's device within space (it's "geopose") and to convey a holistic understanding of the scene for augmented reality systems to place virtual scenes composed of virtual objects, virtual interfaces and information.

The two primary goals of the OARC Reality Modeling and Mapping Working Group are:

1. To describe the techniques used to map and model reality
2. To provide guidelines for building a standardized set of methods to interface with the data generated using these techniques

Below, we discuss the current techniques for capturing the state of reality known today throughout the industry and academia. It is our goal that we cover most, if not all, possibilities and contributions to this field as it applies to both indoor and outdoor mapping techniques.

It is important to note that processing for many of these techniques can take place either in the cloud or on-device.

The maps that devices capture for AR Cloud purposes are highly localized and decentralized, both in terms of their origins and the location where they are of highest value to users.

Traditional cloud architecture and infrastructure is centralized in large and often remote data centers. This results in a technical mismatch that introduces both high data-transfer costs and hard physical limits of latency performance when accessing the data. Ideally, most said data should be stored physically near its origin.

This provides a great case for so-called “edge computing”, a subset of cloud computing where either devices (device edge) or smaller data centers close to wireless data transfer points such as WIFI, 4G, and increasingly 5G (infrastructure edge) do the computing.

For time-critical reality modelling, such as determining real-time states of vehicles and people in traffic and other real-time semantic scene understanding use-cases, edge computing might be the only possible way to deliver it with an acceptable latency.

See the section *Edge Computing and IoT for AR Cloud*” for a more in-depth exploration of this promising technology.

Another promising approach to decentralization of both storage and computation is via distributed ledger technology, such as blockchains. The ability to embed automated contracts and high levels of data security could help solve some very hard challenges in relation to digital rights, privacy, intellectual property and more. See the section *Distributed Ledger Technology and the Blockchain* for more information on decentralized apps for AR Cloud.

Finally, the sheer level of capital expenditure for a centralized storage of intricate maps of the entire world risks being astronomical at current costs of data storage, making such an endeavour something that only tech giants and governments of large countries could hope to undertake today. This would entail that only such organizations could make use of this data at that scale thus creating an unhealthy concentration of power.

Privacy and Security

Privacy is a guiding principle of Open AR Cloud and will play a very important role in defining how to map real spaces, how to store those maps, and ultimately who will have access to them.

As this WG works through the challenges faced by mapping and modeling reality, it will need to carefully incorporate the work that the *Privacy and Security* and *Compliance and Regulations* WGs are doing. It is critical that the systems built for capturing and storing these maps of real spaces have the privacy concerns of that space and its owners at the forefront, and that they comply with local regulations for the management of such data.

Key Enablers

In this section, we take an in-depth look at key technologies that enables reality capture and modeling:

Sensors for Reality Capture

The first step towards creating indexed machine readable maps of the world starts with machines that can capture data about its environment using sensors. The focus in the AR-cloud industry is camera-based reality capture, but in other fields of spatial computing (such as autonomous vehicles and robotics), deploy a wider range of sensors both for redundancy and to capture reality under a variety of conditions. Each sensor type has its own strengths and weaknesses.

This section's authors are of the opinion it is unlikely that AR devices of the future will work optimally in all conditions if they only rely on camera-based sensors .

The table below provides a comparison of sensors used in mapping and localization. Different sensors have different operating characteristics in terms of weather, illumination, resolution, and range. Size and cost are additional considerations for an application. There is no perfect sensor for every situation. For example, delivery robots and self-driving cars typically combine different complementary sensors.

Sensor	Blinding Sunlight	Darkness	Rain, Fog, Snow	Non-metal Objects	Resolution	Range	Size (small = 1)	Cost (low = 1)
RGB Camera	no	no	no	yes	+++	+++	+	+
Event Camera	yes	yes	no	yes	+	+++	+	+++
Infrared	no	yes	no	yes	-	+	+	+
LIDAR	yes	yes	no	yes	++	++	+++	+++
Radar	yes	yes	yes	no	+	+++	+	++
Ultrasonic	yes	yes	yes	yes	+	+	+	+

Sensor Comparison

Likewise, if the real world spatial web will be accessible all day long, in all weather and light conditions, indoors and outdoors, either the devices must be equipped to sense environments given such variations, or environmental-based sensors aid in the process of real-time reality capture. Also, AR Cloud solutions that support fundamental concepts like localization and mapping would benefit from being able to handle complementary types of sensor data.

Current AR techniques that rely solely on cameras for reality capture are severely limited. Until this improves, it is likely that users will be reluctant to make the move to all-day-long wearable AR glasses.
Going forward, we recommend that the AR Cloud industry mimics the autonomous vehicle industry by striving to make systems that handle a sufficient range of complementary sensor input, so they work for users in as wide a range of conditions as possible.

Visual Mapping (For Spatial Mapping)

Mapping Vehicle. Source: Here 360[7]

Highly accurate visual localization drives a need for highly accurate 3D models and maps. Today, mapping providers generate high-definition maps by using professional-grade cameras coupled with other professional-grade sensors such asLIDAR, radar, IMU, and GPS typically attached to vehicles, backpacks or trolleys. Captured data is stored locally, then uploaded and processed offline, often with some manual intervention. This process is costly and slow, and map data can quickly become stale. Ideally,open crowdsourcing with consumer-grade devices will generate accurate models and maps in a fast, fully automated pipeline. With continuous updates. This is a difficult challenge that needs solving, yet it is an important prerequisite for an Open AR Cloud.

Camera Types

Standard consumer cameras normally fall into the categories of *monocular RGB*, *stereo RGB*, or depth sensing *RGB-D*. Most smartphone cameras are monocular, and unable to provide depth data. Some smartphones provide two cameras, and some dual camera configurations can operate as a stereo pair. As an example, the Apple iPhone X includes two rear cameras, one with a telephoto lens and the other a wide-angle lens[8]. The iOS SDK can perform disparity calculations and provide depth data for the stereo pair[9]. Specialized stereo cameras are available for robotics and AR/VR

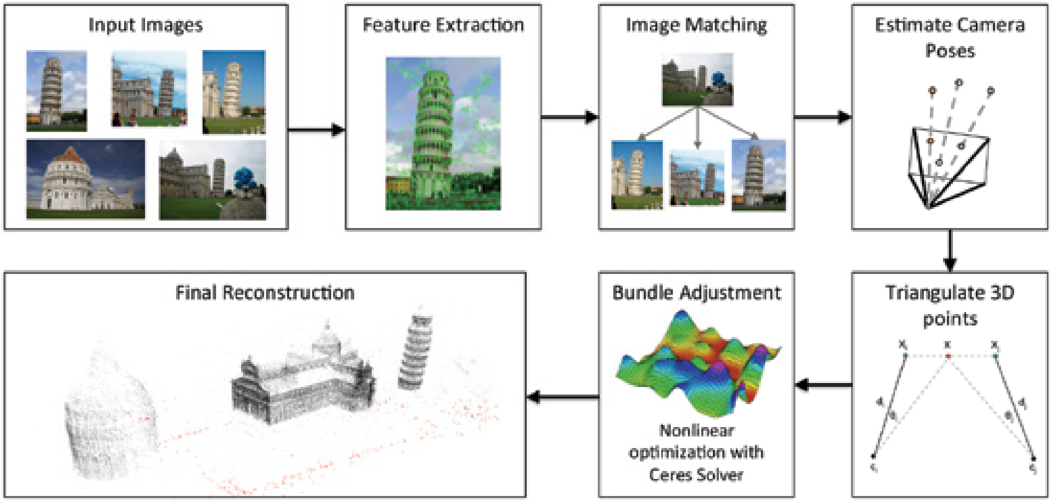
---

[7] Poor, Gwennie, Anatomy of Mapping Cars – for Dummies (April 7, 2015), https://360.here.com/hs-fs/hubfs/Imported_Blog_Media/TrueMappingCar-2.jpg?width=1024&name=TrueMappingCar-2.jpg.
[8] Apple., IPhone XS - Technical Specifications, https://www.apple.com/iphone-xs/specs/.
[9] Apple., Apple Developer Documentation, *https://developer.apple.com/documentation/avfoundation/avdepthdata*.

applications, such as Stereolabs ZED[10] and Intel Realsense T265[11]. Stereo depth estimation works indoors and outdoors, but the accuracy drops off substantially with distance. Active depth sensing RGB-D cameras normally use infrared (IR) technologies. For instance, the front camera on an Apple iPhone X uses IR for accurate facial recognition. Specialized RGB-D cameras are also available, including the Azure Kinect[12], the Intel Realsense D435i[13], and the Occipital Structure Sensor[14]. IR depth sensors typically provide high accuracy over shorter ranges and work best indoors. Robotics applications that require long-range accurate depth sensing in outdoor scenarios (such as autonomous vehicles) rely on LIDAR or radar sensors. Event cameras and, Dynamic Vision Sensors (DVS) provide an innovative vision system for rapidly moving robots, like drones. A DVS transmits per-pixel data only for pixels that change, rather than full frames. Besides gaining a much higher temporal resolution, event cameras are smaller, more power efficient, and provide a higher dynamic range than frame based cameras.[15]

Three key technologies for visual mapping include: Structure-from-Motion (SfM), Visual Simultaneous Localization And Mapping (VSLAM), and Multi-View Stereo (MVS).

Structure-from-Motion (SfM)



Structure from Motion Pipeline. Source: Theia Vision Library[16]

[10] Stereo Labs, ZED Stereo Camera, https://www.stereolabs.com/zed/.
[11] Intel, Tracking Camera T265 – Intel RealSense Depth and Tracking Cameras, https://www.intelrealsense.com/tracking-camera-t265/
[12] Microsoft, Azure Kinect DK, https://azure.microsoft.com/en-us/services/kinect-dk/
[13] Intel, Depth Camera D435i https://www.intelrealsense.com/depth-camera-d435i/
[14] Occipital, Structure Sensor, https://structure.io/
[15] Gallego et al., Event-based Vision: A Survey, http://rpg.ifi.uzh.ch/docs/EventVisionSurvey.pdf
[16] Sweeney, Chris, Theia Vision Library, http://theia-sfm.org/

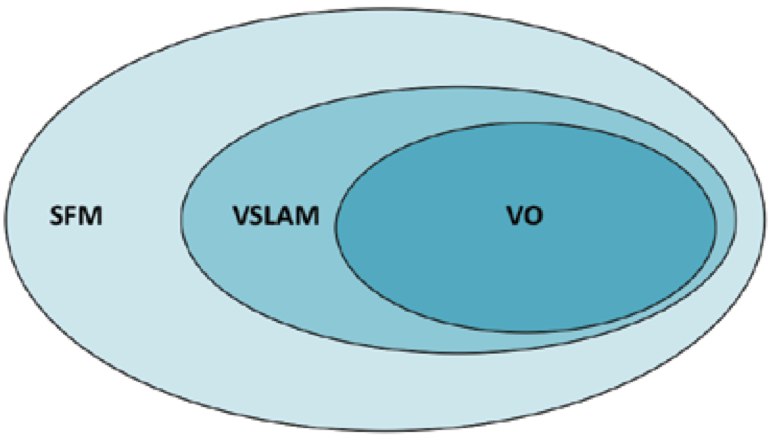
SfM is related to the old art of photogrammetry, where manual measurements from photographs would allow creating a map or model of a scene. SfM solves the problem in an automated pipeline and early work on SfM was performed by Snavely et al. in 2008[17]. SfM creates a sparse 3D point cloud from an unordered set of images. The pipeline extracts image features, matches them, estimates camera poses, and creates a reconstruction by triangulating the points. SfM pipelines usually create high quality point clouds via slow offline processing. There are several strategies for SfM (global, incremental, hierarchical), These strategies can be improved by integrating additional sensor data (ex. GPS, IMU) and GPUs, which can improve performance.

For the AR Cloud, the SfM pipeline would primarily reside off-board (server-side), and process images received from the on-board (client-side) component in a collaborative mapping framework capable of requesting map generation for missing areas. Poiesi et al illustrates this concept[18], critical in addressing a common issue in SfM pipelines where some images fail to register.

Some open source (and often "research-grade") implementations of SfM include:OpenSfM[19], COLMAP[20], Theia, and OpenMVG[21].

[17] Snavely, Noah, et al, Modeling the World from Internet Photo Collections, International Journal of Computer Vision, vol. 80, no. 2, 2008, pp. 189–210.
[18] Poiesi, Fabio, et al, Cloud-Based Collaborative 3D Reconstruction Using Smartphones, Proceedings of the 14th European Conference on Visual Media Production (CVMP 2017), ACM, 2017, p. 1.
[19] Mapillary, Open Source Structure From Motion Pipeline, https://github.com/mapillary/OpenSfM
[20] COLMAP, COLMAP 3.6 Documentation, https://colmap.github.io/
[21] OpenMVG, *Open Multiple View Geometry Library,* https://github.com/openMVG/openMVG

Visual Simultaneous Localization And Mapping (VSLAM)



SFM vs VSLAM vs VO. Source: Davide Scaramuzza[22]

VSLAM and SfM are close relatives. Robotics developers use VSLAM applications where additional constraints apply. Specifically, VSLAM works in real-time with an ordered sequence of images (i.e., a video). VSLAM uses data from monocular, stereo ("passive" depth), or RGB-D ("active" depth) cameras. This may fuse with data gathered from accelerometers and gyroscopes within IMUs. VSLAM usually runs on-board and creates lower quality point clouds than SfM. VSLAM mechanisms can use feature-based mapping to create sparse point clouds, or direct image mapping to create more dense point clouds. Sparse models limit the extent to which AR content

[22] Scaramuzza, Davide, Visual Odometry and SLAM: Past, Present, and the Robust-Perception Age, 2016, p. 127

can interact with a scene. For example, real objects cannot occlude virtual objects with too sparse a model.

An Open AR Cloud could leverage both off-board SfM and on-board (or hybrid on-board/off-board) VSLAM for aspects of visual mapping. The potential low latency associated with edge cloud and 5G provide new opportunities for off-boarding components of a typical VSLAM pipeline.

Challenges for VSLAM include: pure rotation, map initialization, estimating intrinsic camera properties, rolling shutter distortion, and scale ambiguity[23]. Depth data from RGB-D cameras, stereo cameras, or deep learning estimation is valuable in addressing the challenges of pure rotation and scale ambiguity. Fusing IMU data also helps with scale ambiguity. In terms of intrinsic camera properties, it is important to maintain a database of camera properties for specific devices and to identify the devices automatically. Rolling shutter is a common approach in consumer cameras. Each row of an image comes from a different pose, forcing this to factor into camera pose estimation.

Some mobile AR SDKs, such as ARCore[24] and ARKit[25], implement a subset of VSLAM known as Visual Inertial Odometry (VIO). VIO focuses on estimating current 3D motion only, without creating a global map, or performing global localization. VIO can significantly drift over time. ARCore and ARKit also create very sparse point clouds, which limits their mapping capabilities.

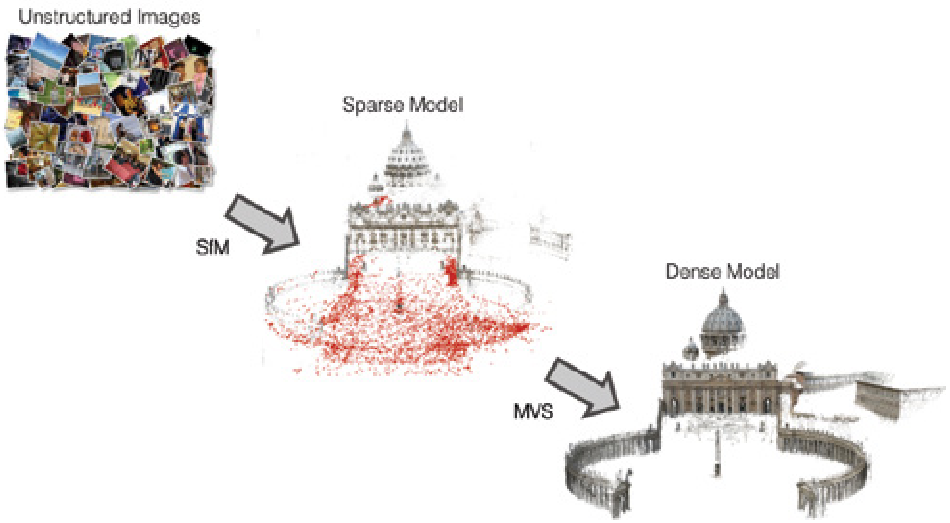
Some open source (and often "research-grade") implementations of VSLAM with mapping and loop closure include ORB-SLAM2[26] and LSD-SLAM[27]. ORB-SLAM2 uses indirect (feature-based) parameter estimation and creates a sparse representation, whereas LSD-SLAM uses direct (photometric) parameter estimation and creates a semi-dense representation

Multi-View Stereo (MVS)

SfM produces sparse point clouds. As with sparse VSLAM systems, this creates issues for AR content interaction. MVS reconstructs scene surfaces of sparse models and creates dense models.

[23] Taketomi, Takafumi, et al., Visual SLAM Algorithms: A Survey from 2010 to 2016, IPSJ Transactions on Computer Vision and Applications, vol. 9, no. 1, 2017, p. 16.
[24] Google, ARCore, https://developers.google.com/ar/
[25] Apple, ARKit, https://developer.apple.com/arkit/
[26] Myr-Artal, Raul, Real-Time SLAM for Monocular, Stereo and RGB-D Cameras, with Loop Detection and Relocalization Capabilities, October 2017, *https://github.com/raulmur/ORB_SLAM2*
[27] TUM Computer Vision Group, LSD-SLAM, December 2014, *https://github.com/tum-vision/lsd_slam*

MVS processing can be slow at scale. Unstructured data further compounds the challenge. MVS in an AR Cloud environment needs the flexibility to deal with both structured (i.e, single-device video) and unstructured data (i.e., overlapping areas or multiple devices). MVS algorithms use different mechanisms to improve performance. This includes intelligently selecting the best subset of images for a cluster, ensuring that operations can parallelize[28], and estimating and merging depth maps[29].



3D Reconstruction Pipeline Source: Furukawa

Some open source (and often "research-grade") implementations of MVS include OpenMVS[30] and COLMAP[31]

Deep Learning

While deep learning has not replaced computer vision in overall VSLAM and SfM pipelines, it is able to address some specific sub-tasks.

Depth estimation can occur via supervised, unsupervised, or semi-supervised methods. For example, Monodepth, an unsupervised mechanism, performs depth

[28] Furukawa, Yasutaka, et al., Towards Internet-Scale Multi-View Stereo, 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2010, pp. 1434–1441.
[29] Middelberg, Sven, et al., Scalable 6-Dof Localization on Mobile Devices, European Conference on Computer Vision, Springer, 2014, pp. 268–283.
[30] cDc, Open Multi-View Stereo Reconstruction Library, 2015, https://github.com/cdcseacave/openMVS
[31] COLMAP, COLMAP 3.6 Documentation, https://colmap.github.io/

estimation from a single input image[32]. The Convolutional Neural Network (CNN) learns the disparity required to map a left-image to a right-image in a stereo pair and vice versa.

CNNs such as FlowNet2 can estimate optical flow[33], and other CNNs including NetVLAD [dl-3] can perform visual place recognition for loop closure [34].

3D Representations

A number of different formats can represent 3D models, including point clouds, voxel grids, octrees, and signed distance functions (SDFs)[35]. Point clouds use unordered data without discretization. There are no limits on the area mapped, but memory consumption explodes. Voxel grids use a more efficient volumetric representation, but with area estimates made ahead of time, and memory already pre-allocated. Discretization results in information loss and creates the potential for processing errors. As with voxel grids, octrees use probabilistic updates, but are multi-resolution and memory efficient. SDFs measure the distance of voxels to the surface to generate highly accurate models. Computing in parallel for all voxels allows for fast operation on GPUs.

Point Cloud[29] *TSDF (KinectFusion)[36]* *Octree[29]*

Visualizations of 3D Representations. Sources: Wolfram Burgard, Nic Fleming

[32] Godard, Clément, et al., Unsupervised Monocular Depth Estimation with Left-Right Consistency, September 2016, http://arxiv.org/abs/1609.03677 .
[33] Ilg, Eddy, et al., FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks, Decemeber 2016, http://arxiv.org/abs/1612.01925.
[34] Arandjelović, Relja, et al., NetVLAD: CNN Architecture for Weakly Supervised Place Recognition, ArXiv:1511.07247 [Cs], Nov. 2015. arXiv.org, http://arxiv.org/abs/1511.07247.
[35] Burgard, Wolfram, et al., Introduction to Mobile Robotics, 2018, p. 46
[36] Fleming, Nic,, High Fidelity 3-D Images Created Using KinectFusion, MIT Technology Review, https://www.technologyreview.com/s/425596/high-fidelity-3-d-images-created-using-kinectfusion/

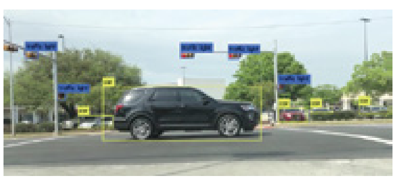
Scene Understanding

Scene understanding is about recognizing an environment and its contents. It includes the spatial and semantic relationships of objects in the environment. Scene understanding in the AR Cloud directly applies to reality mapping and localization, providing valuable insights to higher application layers.

A major challenge for reality mapping is dealing with dynamic scenes. This may involve moving objects like people or cars, or stationary objects that could move like a person sitting down, or a parked car. Ideally, this knowledge factors into the mapping of the static aspects of the environment. It is desirable to track characteristics of moving objects, their poses, and trajectories in real-time.

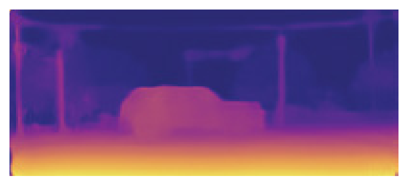
Once an existing reality map incorporates scene understanding, it may match that data against current scene understanding data (assuming real-time processing) for more accurate localization.

Deep learning approaches heavily dominate scene understanding techniques, which involve lower level primitives such as object detection and semantic segmentation.



Object Detection
YOLO v3

Semantic Segmentation
ICNet



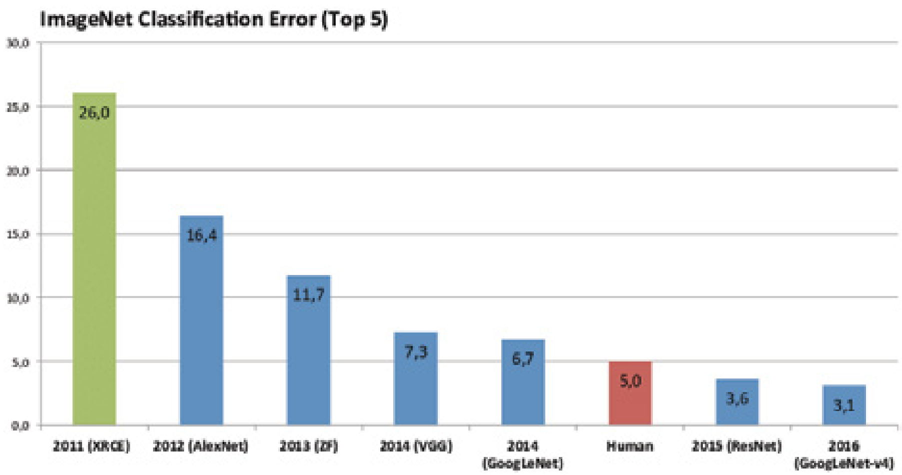
Monocular Depth
Estimation
Monodepth

Scene Understanding Examples.

2D/3D Object Detection

Object detection classifies the objects within an environment and determines their location. CNNs have proven very successful with supervised approaches where models train on a labelled set of objects. Object detection can challenge for a variety of reasons including cluttered scenes, occlusion, deformation, illumination, and shadows.

ImageNet has played a significant role in the recent success of object detection[37]. ImageNet consists of 14+ million images across over 21,000 categories. The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) has seen classification error drop from 26% in 2011 to a remarkable 2.3% in 2018, with human error being around 5%.



Source: ImageNet

Systems exist for detecting objects both in 2D and 3D. Today, 2D mechanisms for object detection are popular, and perform well. The top performing 2D object detection models on the KITTI driving benchmark have an average precision of over 91% for cars and 80% for pedestrians[38] . There are two-stage approaches that involve creating bounding boxes and subsequently performing classification e.g., R-CNN[39], Fast R-CNN[40], and Faster R-CNN[41]. There are also single-stage approaches that directly regress to locations and classifications, e.g., YOLO[42], and SSD[43].

3D object detection is less mature than 2D, but highly relevant to the AR Cloud. The top performing 3D object detection models on the KITTI driving benchmark[44] have

[37] ImageNet. http://www.image-net.org/
[38] Geiger, Andreas et al., Object Detection Evaluation 2012, The KITTI Vision Benchmark Suite, 2012, http://www.cvlibs.net/datasets/kitti/eval_object.php?obj_benchmark=2d
[39] Girshick, Ross, Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, November 2013, http://arxiv.org/abs/1311.2524.
[40] Girshick, Ross, Fast R-CNN, April 2015, http://arxiv.org/abs/1504.08083.
[41] Ren, Shaoqing, et al., Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, June 2015, http://arxiv.org/abs/1506.01497
[42] Redmon, Joseph, et al., You Only Look Once: Unified, Real-Time Object Detection, June 2015, http://arxiv.org/abs/1506.02640.
[43] Liu, Wei, et al., SSD: Single Shot MultiBox Detector, vol. 9905, 2016, pp. 21–37. http://arXiv.org,
[44] Geiger, Andreas, et al., 3D Object Detection Evaluation 2017, The KITTI Vision Benchmark Suite, 2017, http://www.cvlibs.net/datasets/kitti/eval_object.php?obj_benchmark=3d

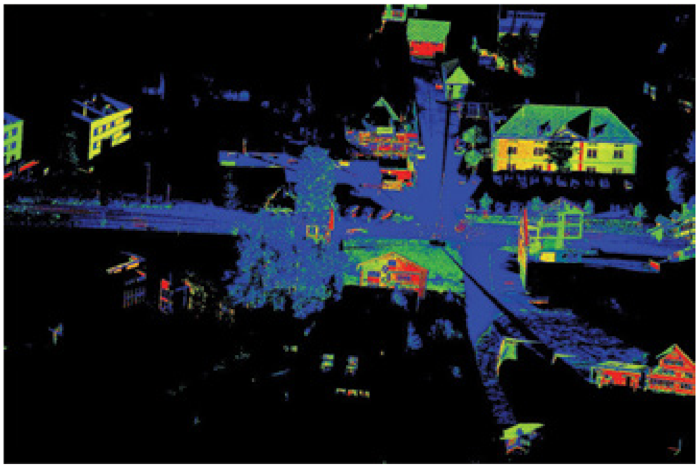
average orientation similarities of only 77% for cars and 46% for pedestrians . 3D object detection can work on front-view images with depth data, bird's eye views, or 3D point clouds. State-of-the-art networks such as Frustum Pointnets[45] leverage both 2D image data and depth from 3D point clouds by adapting PointNets[46] PointNets extracts features directly and describes them globally, Before combining the point and global representations for final predictions. Outside of PointNets, some mechanisms add structure to a 3D point cloud (e.g., voxelization) and subsequently apply CNNs. However, voxelization approaches create additional overhead and result in loss of information.

Semantic Segmentation (2D/3D)

Semantic segmentation provides a pixel-by-pixel assignment to specific categories. Scene complexity and the number of categories present challenges for semantic segmentation. As with object detection, CNNs have proven very successful for semantic segmentation. Some early networks for semantic segmentation (circa 2014 - 2015) were fully convolutional neural networks (FCNNs)[47] and encoder-decoder networks such as SegNet[48]. Popular datasets for semantic segmentation include: PASCAL VOC2012[49], Cityscapes[50], MS COCO[51], and ADE20K[52]. State-of-the-art networks on Cityscapes with open source code include DPC[53], SSMA, and DeepLabv3+[54].

[45] Qi, Charles R. et al., Frustum PointNets for 3D Object Detection from RGB-D Data, November 2017, http://arxiv.org/abs/1711.08488.
[46] Qi, Charles .R et al., Pointnet: Deep Learning on Point Sets for 3d Classification and Segmentation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 652–660.
[47] Long, Jonathan, et al., Fully Convolutional Networks for Semantic Segmentation, November 2014, http://arxiv.org/abs/1411.4038.
[48] Badrinarayanan, Vijay, et al., SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, November 2015, http://arxiv.org/abs/1511.00561.
[49] Everingham, Mark, The PASCAL Visual Object Classes Challenge 2012, http://host.robots.ox.ac.uk/pascal/VOC/voc2012/
[50] Cityscapes Dataset, Semantic Understanding of Urban Street Scenes, https://www.cityscapes-dataset.com/
[51] COCO, Common Objects in Context, http://cocodataset.org/#home
[52] ADE20K Dataset, http://groups.csail.mit.edu/vision/datasets/ADE20K/.
[53] Chen, Liang-Chiehet al., Searching for Efficient Multi-Scale Architectures for Dense Image Prediction, September 2018, http://arxiv.org/abs/1809.04184.
[54] Valada, Abhinav, et al., Self-Supervised Model Adaptation for Multimodal Semantic Segmentation, August 2018, http://arxiv.org/abs/1808.03833.

3D semantic segmentation is less mature than its 2D counterpart. Semantic3D.net[55] provides a large-scale point cloud classification benchmark. PointNets[56] also provide cutting edge performance for 3D semantic segmentation.



Semantic Segmentation on 3D Point Cloud. Source: Semantic3d.net

Temporal Aspects

The temporal components of a scene support operations such as object tracking and activity detection, providing valuable input for object detection and semantic segmentation. For example, person detection can leverage leg movement as an input.

Optical flow measures the 2D motion of brightness patterns between a pair of images. Sintel[57] and the KITTI optical flow[58] datasets provide benchmarks for optical flow. CNNs such as FlowNet2[59] demonstrate state-of-the-art performance for optical flow.

---

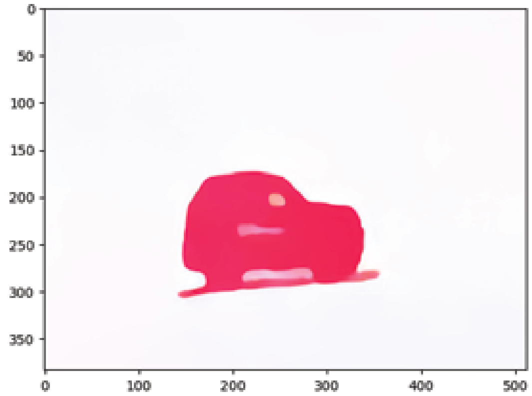
[55] http://www.semantic3d.net/

[56] Qi, Charles .R et al., Pointnet: Deep Learning on Point Sets for 3d Classification and Segmentation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 652–660.

[57] Max Planck Institute, MPI Sintel Dataset, http://sintel.is.tue.mpg.de/.

[58] Geiger, Andreas et al., Scene Flow Evaluation 2015, The KITTI Vision Benchmark Suite, http://www.cvlibs.net/datasets/kitti/eval_scene_flow.php?benchmark=flow

[59] Ilg, Eddy, et al., FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks, December. 2016, http://arxiv.org/abs/1612.01925.



Optical Flow. Source: FlowNet2

Scene flow adapts optical flow to a 3D environment. At a minimum, scene flow relies on multiple stereo images of a scene. The KITTI scene flow[60] dataset provides a key benchmark. PRSM[61] is a leading network for scene flow with open source code.

Object tracking allows for object correlation across images in a sequence, and motion forecasting predicts where the objects will be in their near future. As examples, Tao et al.[62] perform 2D tracking via siamese matching networks and Alahi et al.[63] predict future human trajectories by using a long short-term memory (LSTM) network. Luo et al.[64] use a neural network to reason about 3D detection, tracking, and motion forecasting.

Beyond Basic Scene Understanding

Based on the techniques described, the reality model in the AR Cloud could consist of a rich 3D model with a semantic segmentation overlay, and static or dynamic 3D object overlays. While many challenges exist in realizing such a model using crowd-sourced consumer devices, this is only the foundation of scene understanding. It is ultimately desirable to evolve scene intelligence to incorporate context and relationships.

---

[60] Geiger, Andreas et al., Scene Flow Evaluation 2015, The KITTI Vision Benchmark Suite, http://www.cvlibs.net/datasets/kitti/eval_scene_flow.php?benchmark=flow

[61] Vogel, Christoph, et al., 3D Scene Flow Estimation with a Piecewise Rigid Scene Model, International Journal of Computer Vision, vol. 115, no. 1, 2015, pp. 1–28.

[62] Tao, Ran, et al. Siamese Instance Search for Tracking, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1420–1429.

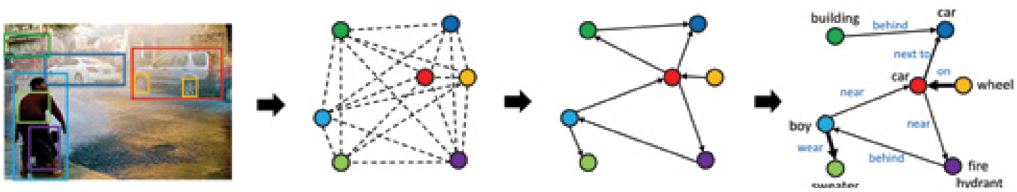
[63] Alahi, Alexandre, et al., Social Lstm: Human Trajectory Prediction in Crowded Spaces, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 961–971.

[64] Luo, Wenjie, et al., Fast and Furious: Real Time End-to-End 3d Detection, Tracking and Motion Forecasting with a Single Convolutional Net, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3569–3577.

Semantic graphs are a common method for describing the loose relationships between entities contained within a domain. For example, the World Wide Web Consortium has published a semantic graph standard for reliably sharing content across the web called the Resource Description Framework [65]. This standard allows loosely coupled independent data structures to be shared across owners and time.

The AR Cloud could include a semantic spatial graph system, capable of capturing and maintaining the constantly changing state of objects contained within a real space, for the purpose of understanding by a computer system. This *reality graph* would become a standard for reality mapping and reference in the use case of AR applications.

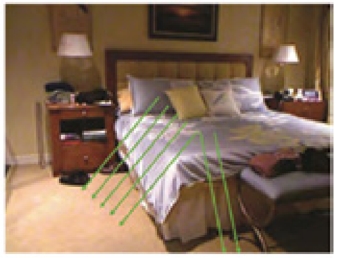
Object detection models could directly output the reality graph. For example, Yang et al.[66] propose a scene graph generation model, Graph R-CNN, that detects objects and their relations in images.



Stages of Graph R-CNN. Source: Jianwei Yang

Per Naseer et al.[67], physics-based reasoning includes estimating current and future dynamics from a static scene, understanding the support relationships and stability of objects, volumetric and occlusion reasoning. Affordance prediction goes beyond context and relationships of objects to understanding their functionality and affordances, i.e., that one may "sit" on a chair. Activity recognition provides visual reasoning around the activities or actions within a scene. This includes specific scenarios such as gesture recognition, and may leverage input from special models for human pose estimation. Activity recognition also enables affordance prediction i.e., the reasons why "sitting" is associated with chairs. Physics-based reasoning, affordance prediction, and activity recognition provide valuable input around placing and interacting with virtual content in the real world.

[65] W3C, RDF - Semantic Web Standards, https://www.w3.org/RDF/
[66] Yang, Jianwei, et al., Graph R-Cnn for Scene Graph Generation, Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 670–685.
[67] Naseer, Muzammal, et al., Indoor Scene Understanding in 2.5/3D for Autonomous Agents: A Survey, IEEE Access, vol. 7, 2019, pp. 1859–1887



Physics-based Reasoning. Source: Muzammal Nassser

Affordance Prediction. Source: Muzammal Nassser

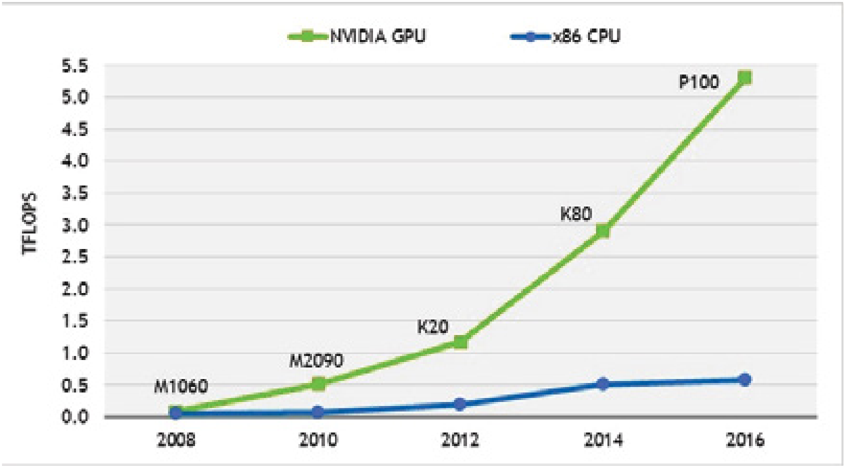
A Universal Scene Description (USD) Protocol

AR Cloud entities must ultimately communicate scene understanding data in real-time. This requires an efficient, standardized, Universal Scene Description (USD) protocol. The USD protocol would ideally encode the "reality graph". The XVIZ protocol[68], developed by Uber, is an open protocol that provides a stream-oriented view of a scene changing over a period of time. XVIZ describes visual elements including geometry, point clouds, images, text, and metrics.

Model Management and Performance

Given the prevalence of deep learning in scene understanding, managing models becomes a key consideration. This includes aspects such as model training (updates) and model distribution. Model training is typically an offline process where a new model is trained, tested, and subsequently activated. Depending on the specific scenarios, model prediction may operate in real-time or offline. Real-time model prediction may occur on-board or off-board, and different models may dynamically activate in different scenes. For example, if a user moves from an outdoor location to an indoor location, scene understanding could shift to using models optimized for indoor environments. Finally, the parallelism of GPUs and their generational performance improvements, as shown below, has played a very important role in the recent success of deep learning.

[68] AVS, XVIZ - A Protocol for Real-Time Transfer and Visualization of Autonomy Data, https://avs.auto/#/xviz/overview/introduction



Source: NVIDIA

Challenges and Opportunities

Looking ahead to a future that features a well-functioning AR Cloud, the main challenges and opportunities that face Open AR Cloud and the rest of the augmented reality industry, follow three themes: interoperability, standardization and scale. Scale and interoperability largely depend on industry standardization. If industry standards are applied to the collection and processing of spatial data, then most techniques for modeling reality should be useful and deployable for an AR Cloud.

Interoperability

Interoperability is fundamental to a shared augmented reality experience. No matter the methods used to build a map or understanding of a real space, that map should work across all modes of viewing and interaction. This level of interoperability is achievable through adopting a set of industry-approved standards.

Standards

Consider the standards that apply to the web regarding content delivered in standardized ways. It matters less how to generate and store content, as long as it follows the standards implemented in the browser of choice when served to the end-user.

To address these challenges, we plan to further explore in collaboration with the industry what set of new standards might be necessary, or what existing standards might necessary to extend, for creating spatial maps and semantic scene graphs that are interoperable across a variety of platforms, devices, and capture techniques.

Existing international standards exist for delivery of 3D data to the browser. Examples known to work in a geospatial context include the Web3D Consortium's X3D[69] (also ratified by ISO) and the Khronos Group's glTF[70]. Standards also help act as the intermediary between the 3D content models and their eventual delivery to the browser. The OGC has three approved standards for 3D scene composition for delivery: the OGC 3D Portrayal Service standard[71] (3DPS), the Indexed 3D Scene Layer and Scene Layer Package Format Community standard[72] (I3S); and the 3D Tiles Community standard[73].

Scale

It is not lost that the scale of mapping and modeling reality is beyond nearly anything humankind has attempted to date. This is why industry standards must be applied. Otherwise, we will end up with disparate, quasi-functional implementations that work only in specific use-cases and only with specific interfaces.

Proper scaling means that we need to move away from purely proprietary data structures for storing SLAM data and consider an open, standardized approach that the industry can contribute to.

Freshness of data

Location data, whether aggregated automatically or manually, changes continuously. A simple example is traffic data, which affects the value of routing and location-related services and depends highly on being as up-to-date as possible. The same reasoning applies to all data attributed to time and space and shared by more than a single user within the AR cloud.

On top of standardization, the technical challenges of scales of the magnitude proposed by mapping and modeling reality, are great. The technical community will need to collaborate to realize the concepts of an interoperable world-scale system such as proposed by the AR Cloud.

[69] http://www.web3d.org/standards/number/19775-1
[70] https://www.khronos.org/gltf/
[71] https://www.opengeospatial.org/standards/3dp
[72] https://www.opengeospatial.org/standards/i3s
[73] https://www.opengeospatial.org/standards/3DTiles

Spatial Indexing and Geopose

Introduction

In this section, the focus is on technologies and tools which will address several fundamental gaps in the current infrastructure to deploy and use AR Cloud and spatial computing:

- The absence of spatial index schema or protocols to populate and maintain spatial registries. When available, experience publishers can use spatial registries to anchor experiences to geospatially unique objects and places, and by their audiences (the consumers of AR experiences) when querying AR Cloud services.
- The absence of protocols and conventions that would permit any user's AR display to use sensors, data and services to obtain its unique pose in 6 dimensions, with respect to the physical world.
- The absence of a shared frame of reference to express, record, and share a universally understood poses of real and virtual objects such as devices, 3d datasets, photos, video frames, vehicles and digital 3D objects.

This section concerns itself with the development of tools, standards and technologies for devices and services to query captured information about the physical world, as described in the section above.

Scope

Anchoring AR content in an AR Cloud architecture requires *spatial indexing* and *geopose* to populate a spatial registry used by associated devices and services. These technologies have synergies with tools for reality capture, protecting privacy , distributing computational resources (across networks and devices), and delivering real time services.

Geopose

Every physical object on Earth inherently has a geographically anchored position and orientation. We can say the same about virtual objects or three-dimensional datasets (such as point clouds) that we would like to anchor to the real world. Unfortunately, to OARC's knowledge there is not any standard for universally expressing the combination of position and orientation in a manner for interpretation and use by modern spatial computing platforms and AR systems.

Definitions of *pose* and *geopose:*

> A real object in space can have three components of translation: up/down (z), left/right (x), and forward/backward (y). It also has three components of rotation: Pitch (θ), Roll (ϕ) and Yaw (ψ) . Hence objects have "six degrees of freedom" (6DOF).
>
> In computer graphics and robotics, the combination of an object's position and orientation is usually referred to as the object's "pose." Pose is expressed in relation to other objects and/or the user. When a pose is defined relative to a geographical frame of reference or coordinate system, we may call it a geographically anchored pose, or "**geopose**" for short.

Several companies in the AR industry (including some founding members of Open AR Cloud, as well as larger companies like Google) are using, developing or delivering visual positioning services that provide geospatially anchored poses of the device itself or the device's local coordinate system. But, they each have their own way of describing such poses and they are not interoperable across their proprietary ecosystems.

While there are many great use-cases for locally anchored poses, it is OARC's opinion that local poses significantly limits what is possible. For a potential future where people wear an AR device most of the day and interact with the digital world through a browser for the "Real World Spatial Web" as a seamless experience, we must enable greater interoperability across devices, content, platforms, solutions and services via a universal geopose standard with a shared geospatial frame of reference.

The current lack of a universal geopose standard is one of the major shortcomings of the AR Cloud industry today. It is holding the industry back from developing a lot of important use-cases for AR Cloud technology.

Benefits of a Geopose Standard:

- A shared frame of reference enables interoperability between different techniques to obtain geopose for AR devices

A single universal frame of reference for geopose makes it entirely possible to represent digital objects in the real world with the same geospatial position and orientation for multiple AR users who may have systems that use very different methods to obtain their device's geopose. Any future method or service that provides geopose in the same frame of reference will provide a consistent result, making the approach both backwards and forwards compatible.

This level of interoperability stands in contrast with anchoring content on or in the physical world using methods that do not share the same frame of reference — much like storing a pose relative to a set of local features is done with shared cloud anchors used by different vendors today (e.g., ARCore Cloud Anchors, ARKit ARWorldMap, Microsoft Azure Spatial Anchors, Magic Leap's Map Merge, etc). With the geopose specification, many users would be able to interact with the same content, using different devices and services. They would not be required to use the same exact method to obtain their pose relative to the local features. This would reduce reliance on a specific approach and reduce the risk of service provider lock-in.

- **Place content relative to the real world declaratively**

A universal frame of reference for declaring a digital object's exact geographical location with an exact orientation will enable many new use cases. This would give any user or system the ability to declare the specific geopose of a digital object, regardless of the geographic location of the user or system.

For example, an[1] architect may place a 3D model of a proposed building at the location they intend to build it, without literally being there. The architect may use digital 3Dmaps of the area to provide assurance they positioned the building correctly. Since the frame of reference is universal, a person at the location could see and move around the designated location and experience the proposed building at a 1:1 scale representation, positioned and oriented exactly as the architect intended.

- **Display digital objects placed relative to a moving frame of reference simultaneously as displaying digital objects placed relative to the stationary world**

Current mobile AR SDKs like ARCore or ARKit cannot display objects relative to a moving frame of reference when a user is inside of that moving frame of reference. If a user is inside a car, train or bus and places digital objects inside the vehicle, the object is likely to drift away as the vehicle moves in ways that cause acceleration or deceleration. However, if the vehicle itself transmitted a stream of geoposes to the devices and services subscribed to the object, the AR software on the device should in theory be able to compensate for those movements in the real world.

An AR system may offer a user the option of placing content relative to the moving vehicle or relative to the world outside. The same principle applies to floors of tall buildings that may move a few feet back and forth during strong winds or earthquakes. The building could have sensors that allow it to provide a real-time data stream of geoposes for the floor and enable AR solutions to account for those movements, both for placing content inside or outside the building.

Another example is for a multi-user application. Attaching digital objects relative to a user as that user moves around will offer compelling game dynamics more familiar in their interactions with the real world. Other users may thus see how those objects “follow” the user’s movements as if being part of that user.

Examples of High Value Industrial Geopose Use Cases

- Construction and Engineering

The industry already uses GIS and geospatial data to a high degree and is increasingly starting to make use of BIM models that personnel can view and interact with onsite using AR-enabled glasses. Standardized geopose would be of great value for this industry reducing the need for manual calibration steps or maybe even the need for expensive, highly accurate GPS receivers on each device.

- Mobility and Transportation

Autonomous vehicles already use services to locate relative to geospatial data to obtain a form of geopose of the vehicle, and to estimate the geopose of vehicles nearby based on processing sensor data with machine learning. With standardized geopose, vehicles can share their geopose with nearby vehicles and pedestrians, even if landscapes, buildings, or other objects obscure them. Transportation companies with fleets of vehicles could use geopose to transmit not only the location of each vehicle, but also to provide services to passengers describing where the vehicle of interest is, and to use live geopose streams to correct for movements to allow correct view of virtual objects positioned both inside and outside of the vehicles.

- Tourism

Companies in this industry could enrich the experience of a tourist site by creating accurately positioned interactive content that draw tourists to the particular site. The content,either free or available for purchase, could display without requiring users to make special configurations. As long as their AR devices have a geopose, and the content has a geopose the experiences should work seamlessly for everyone.

Developing a Geopose Standard

The development of a standard for geopose is one of the key enablers of interoperability in spatial computing (including, but not limited to, the AR Cloud industry), and a fundamental prerequisite for many important use cases. One of the early goals of the OARC is to jumpstart the development of such a standard in collaboration with stakeholders in the spatial computing sector.

Members of OARC have written a proposal for such a “Standards” Working Group, and have submitted it to the Open Geospatial Consortium (OGC), a founding partner of

OARC and the primary standard development organization within spatial computing. The proposed charter is being evaluated to become a Standards Working Group (SWG) within the OGC. The proposal has also been sent to a number of companies and organizations working in the spatial computing sector, including Here Maps, Norkart, The Norwegian Mapping Authority and BuildingSmart International. The initial response is enthusiastic and there is reason to believe that work on developing such a standard could begin in 2019.

If this work succeeds, OARC will have made its first historical contribution to the future of spatial computing, taking a fundamental step to developing a critical component of the scaffolding and glue that will connect the digital and physical worlds.

Spatial Index and Reality Capture

Some of those investing in spatial indexing are combining their technologies with those for reality capture. The spatial index protocols will inherently share many core enablers with reality capture technology. Reality capture systems are populating a digital 3D map or “twin” of the physical world. When indexing a reality map, a database populated with unique keys or features that make them identifiable.

Indexing can take advantage of multiple localization techniques. When available, these can include the capture system and the indexing system’s GPS coordinates. However, relying solely on GPS dramatically constrains the scope and nature of search queries. Spatial indexing to a reality capture process may, in addition to GPS, create a hash on GPS and WiFi SSID to index maps. Physical world addresses can also serve as the basis for indexing a data set. Floors in those addresses can act as an additional component and contribute to a higher level of detail.

Regardless of how many and which techniques systems use during spatial indexing ,to specify a user’s pose in 6D, or to query and use spatial registries, reality capture systems must use syntax and conventions agreed to and implemented in AR displays, content and services.

Private, Shared and Public Access to Location Data

For AR Cloud services, one or more public registries will need to use a schema for pointers to the data anchored to people, places and things in the physical world. Privacy is a fundamental attribute of people and many spaces should not be defined as “public” or “shared”. Therefore, when designing and delivering spatially indexed services or assets, one or more top level registries will need to clearly delineate terms and conditions which will treat the use of public, shared and private registries appropriately.

Geopose is an example of a potential data attribute of one or more devices in use by people in real time. As for the spatial registries, different use cases will define different levels of private, shared or public access. For example, an individual navigating through the real world will seek to have their personally referenced geopose be continuously private to the outside world. However, groups of two or more people in AR experiences, can share their poses with services or other devices. In this case, the access to the location is shared but not necessarily public.

"On-AR" vs "On-Network" Devices

ISpatial indexing and geopose technologies will be fundamental to the delivery of AR experiences consistent with the AR Cloud concept, and to foster the emergence of valuable new services. Like reality capture, spatial indexing, publishing to and by spatial registries, and the use of geopose can occur either on device, on networks , or a combination of these two.

A high-level goal will be to have systems perform these functions, within some constraints, both offline and online. Designers of AR Cloud-capable devices and services will need to provide the ability to store the data necessary for critical computation entirely on device and make it accessible to users in low or zero-connectivity regions.

Real Time Services

In contrast with reality capture, which can be post-processed (hence computationally delayed), querying or "reading" the spatial index and generating a geopose must be real-time functionalities. They are highly time-dependent because, as a user moves in space and time, the results provided by either of these services in the past are no longer "true" in the present.

The same real-time status update capability could also apply to privacy settings of a user in the AR Cloud: in some use cases, the user may wish to share their geopose and status, while in others it may need privacy.

Therefore, processing for both spatial indexing and geopose services must continuously repeat or run in real time for consistent and satisfactory delivery of AR Cloud-enabled experiences.

Key enablers

From a technology perspective there are at least seven key enablers for spatial indexing and geopose: reality capture and mapping; visual localization and positioning; non-visual localization and positioning; high performance servers; high performance, low cost mobile devices; and high bandwidth, low latency networks.

Reality Capture

A fundamental precursor to spatial indexing is the capture of physical world spaces. Before we can generate a spatial index, we must capture physical spaces and their contents. For that, we must construct a technical pipeline for reality capture. It begins as a set of raw data from the environment. While a reality map is being created, there needs to be a method to store the map against a spatial index.

A generic capture pipeline could have these steps:

1. Capture raw data and create the spatial map
2. Segment the scene in 2D (with pixels) or in 3D, (with voxels) either by generic class or unique segments
3. Including other sensor data like WiFi, cell tower, or magnetic data about that point captured.
4. Visual scanning to support identification of unique objects.
5. "Semantic graphing" to describe the relationships between different objects in the reality map

The Need for a Universal Spatial DNS - A Historical Comparison

The internet as we know it today came into existence when Sir Tim Berners-Lee invented the World Wide Web in 1989. Marc Andreesen created web browsers like Mosaic and Netscape, and the whole world started running on the internet. The World Wide Web uses Top Level Domain Name System (DNS) to tell web browsers which computers / IP addresses to connect to when you try to reach a site such as www.google.com. Every time any browser tries to connect to a domain name it hits a DNS Registry first run by companies like Verisign. Sixty percent of top level .net and .com domain names are resolved by Verisign.

History could repeat itself with the emergence of the Spatial Web where data and online experiences will tie-in to the physical world. Persistent AR experiences are at a fraction of their potential without a way to move from one AR experience to the next. Without a global spatial index, finding AR experiences currently is the equivalent of browsing the internet with IP addresses, or needing a new browser for each website.

Transitioning Between Indexes

There are challenges with transitioning between indoor and outdoor localization models, as well as between top level domain and sub-domain indexes. For example, outdoor point clouds based on satellite images or outdoor photography generated by companies like Sturfee and Scape may not be compatible with indoor point clouds generated by 6D.ai.

In its mature implementations, a user's AR Cloud software stack for viewing AR experiences will need to switch seamlessly between one or more spatial index types (or registries) as the user travels in time and space.

Types of Positioning

There are several options which can assist in determining the user's position and pose. For example, in a simple 2D space, the moment a user needs to find a room in a building, visual references might not be available (no street signs indoors). Alternatively, the user might employ the same frame of reference: outside looking in, or inside looking out. In some use cases, users will need to have a shared frame of reference based on absolute position and relative position. The visual cues of going into a room in a similar building are the same in California or China. Hence, a combination of absolute and relative location is necessary to localize users of spatial computing correctly.

The spatial index service could communicate to a subscribed device user about an absolute position in lat / long / height relative to the earth or relative position to a corner in a building untied absolute position. The relative position may also be relevant when there are two similar looking buildings or objects which can be in two or more different absolute locations on earth.

This could lead to object-specific spatial index where each object indexes against a 3D space a query can find which objects are near other objects, and space model indexing where there is a subdivision of the "world" of interest into spaces with some semantic relevance: e.g., rooms in a building, sections in the US public land system, or addresses in a postal system.

A user of spatial indexing may also consider dynamic space modeling, where the spaces definitions are based on the density of objects in each space unit. We lose semantics, but computation is better.

Geospatial Computing for Autonomous Vehicles

That practically all state-of-the-art autonomous vehicle systems with Tesla Motors use so-called HD maps in the cloud to algorithmically match with sensor observations from the vehicle to obtain highly accurate geospatial position and orientation of vehicles, has for a time provided a precursory proof of concept of a geopose-based approach. Similarities with the early AR Cloud visuals positioning methods seems striking but the technological milieu is radically different. AR Devices cannot match the raw processing power and electric power consumption one can put on a car, requiring a different set of algorithms, sensors, and an offloading of compute from device edge to infrastructure edge-compute or cloud-compute.

Visual Localization

Visual localization is a key requirement for accurate 6DOF tracking in the AR Cloud. In particular, resource constrained mobile clients must be able to accurately localize position and orientation against the real-world global map in real-time.

The primary types of localization are: image-based, structure-based, and learning-based.

Image-Based Visual Localization

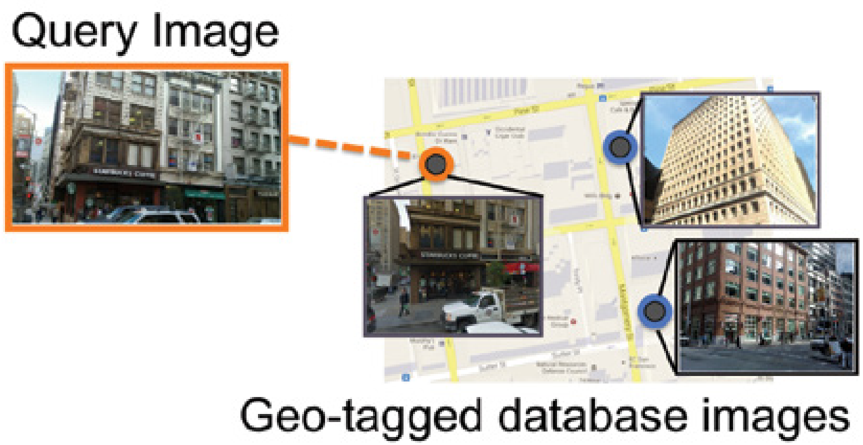
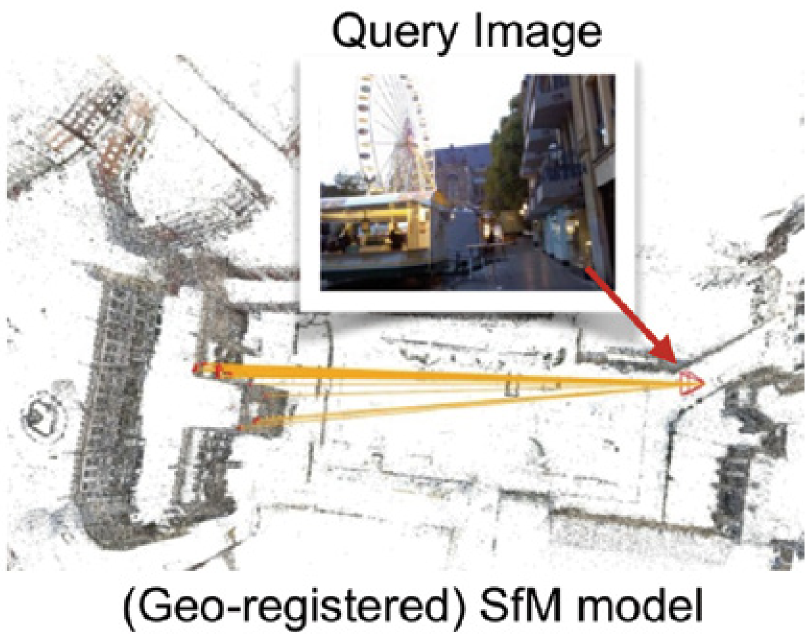


Image-based Visual Localization. Source: Torsten Sadler[74]

Image-based (referring to the database composition) localization relies on computer vision or machine learning approaches to visual place recognition. The process compares images against a known corpus containing associated poses. This provides only a coarse estimate of pose and the accuracy is typically insufficient for AR use cases.

[74] Sadler, Torsten et al., Large-Scale Visual Place Recognition and Image-Based Localization - CVPR 2015, https://sites.google.com/site/lsvpr2015/

Structure-Based Visual Localization



Structure-based visual localization. Source: Torsten Sadler

Structure-based localization involves 2D-3D (or 3D-3D with depth) matching against a 3D model such as a sparse or dense point cloud. SLAM has traditionally provided structure-based real-time localization in robotics applications. VSLAM implementations usually run on-board to ensure real-time performance. In practice, on-board resource constraints limit VSLAM to small areas. These challenges may be addressed by off-boarding some components of VSLAM. For example, Middelberg et al.[75] present an architecture that constrains on-board VSLAM to a small window and performs off-board localization against a global 3D model. This allows the local map to align itself to the global map and avoid drift.

Challenges for structure-based localization include dynamic environments, textureless surfaces, and changing conditions such as illumination, weather, and seasons. With a pre-existing global map, there is an inherent possibility of differences between the global map and the current environment. Appearance-based localization is often more robust than feature-based localization, and including depth data also improves robustness.

[75] Middelberg, Sven, et al., Scalable 6-Dof Localization on Mobile Devices, European Conference on Computer Vision, Springer, 2014, pp. 268–283.

Learning-Based Visual Localization

Various groups have also proposed machine learning approaches to localization. Kendall et al.[76] present a CNN trained with images and their corresponding poses within a specific locale. The network can learn structural properties and predict poses for new (unseen) images within the same locale. The prediction runs in real-time on a GPU, but the accuracy is lower than computer vision approaches and updates require offline retraining of models.

Other Visual Localization Mechanisms

Besides image-based, structure-based, and learning-based localization, users could leverage high-level semantics and cross-view relationships for localization. High-level semantic localization matches constructs such as building models, objects, or semantic segments. Norkart AS proposed on-device detection of semantic object constellations to match with a semantic object model in the cloud in late 2017 for a low budget R&D program still in its infancy. Cross-view localization matches across differing vantages, such as street and aerial views.
Ultimately, hybrid visual localization approaches that combine multiple techniques are likely to provide the most robust results for AR Cloud.

Non-visual localization and positioning

While visual localization is the primary modality for AR Cloud, robust solutions will incorporate localization data provided by non-visual sensors.

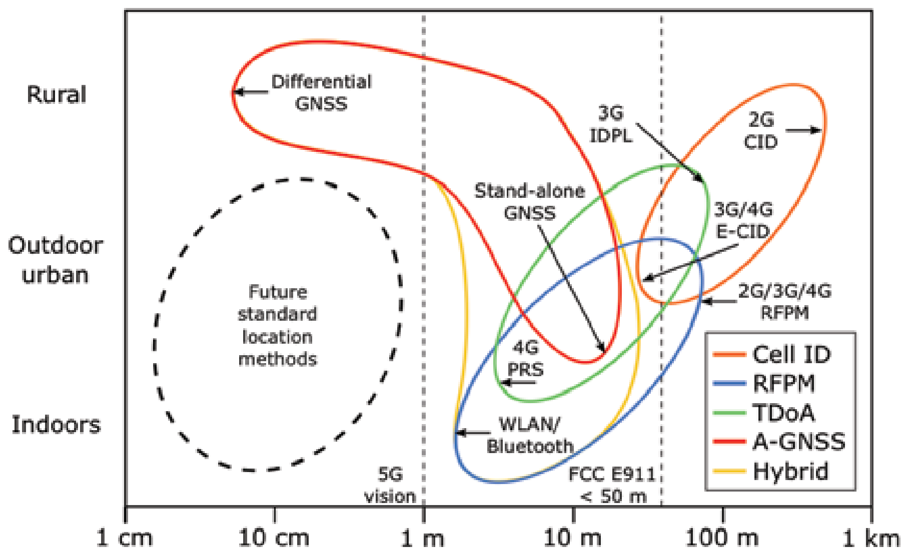
GPS

Smartphone GPS is only accurate to about 5m[77], provides position but not orientation, and requires a clear view of the sky. Smartphone GPS accuracy will increase to tens of centimeters[78] as more chipsets support dual-frequency Global Navigation Satellite Systems (GNSS). Galileo, the new European Union GNSS, provides the additional band. In scenarios where GPS is available, it provides a valuable constraint on the target area for visual localization.

---

[76] Kendall, Alex, et al., Posenet: A Convolutional Network for Real-Time 6-Dof Camera Relocalization, Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 2938–2946.
[77] GPS.Gov: GPS Accuracy. https://www.gps.gov/systems/gps/performance/accuracy/.
[78] IEEE Spectrum, Superaccurate GPS Coming to Smartphones in 2018 .
https://spectrum.ieee.org/semiconductors/design/superaccurate-gps-coming-to-smartphones-in-2018.

Cellular Localization



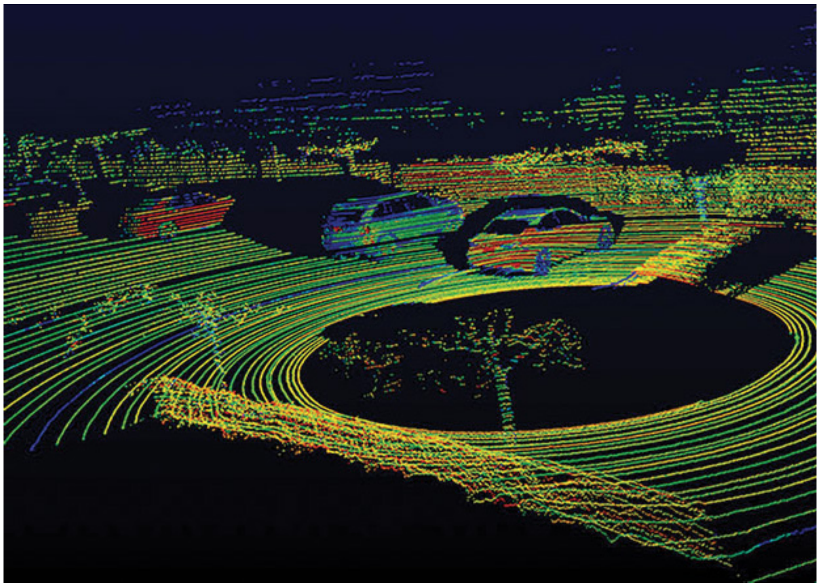
Horizontal accuracy of cellular mobile radio localization methods. Source: José A. del Peral-Rosado

In 4G networks, mechanisms such as time-difference of arrival (TDOA) provide position-only localization with accuracy to tens of meters[79]. As with GPS, this provides a constraint on the visual localization area.

5G standards are still evolving, but with high carrier frequencies in the millimeter-wave (mmWave) band and a large number of antennas, position and orientation can potentially be localized to within 10cm and 1 degree, respectively[80]. As with GPS, this method does not work well indoors, but it works well in the dark, with blinding sunlight, and with fog, snow, or rain. This means 5G localization could be a valuable complement to outdoor visual localization.

[79] del Peral-Rosado, José A., et al., Survey of Cellular Mobile Radio Localization Methods: From 1G to 5G, IEEE Communications Surveys & Tutorials, vol. 20, no. 2, 2018, pp. 1124–1148.
[80] Wymeersch, Henk, et al., 5G MmWave Positioning for Vehicular Networks, IEEE Wireless Communications, vol. 24, no. 6, 2017, pp. 80–86.

Robot Localization



LIDAR Point Cloud. Source:IEEE Spectrum [81]

AR Cloud clients may include robotics applications such as road vehicles, unmanned aerial vehicles (UAVs), and sidewalk delivery robots. In these applications, LIDAR, radar, and ultrasonic sensors could assist with localization.

Other Indoor Localization

Technologies such as WiFi, Bluetooth Low Energy (BLE), ultra-wideband (UWB), or Near Field Communication (NFC) may assist with localization in indoors scenarios..

Computational Resources

To perform the localization and continuously provide 6DOF pose, the user's device must have access to high performance computational resources. These resources may be on-device or off-device or a combination of the two. If the computational resources are local (on device), then they must be low power consuming which, typically, means optimized for the functions (not general purpose computing resources).

Along with on-device computing, there willalso need to be off-device computing resources. These may be anywhere including but not limited to "the cloud." As many realize, the cloud is not one place or thing. It is the assumption of this working group

[81] IEEE Spectrum, Cheap Lidar: The Key to Making Self-Driving Cars Affordable, https://spectrum.ieee.org/transportation/advanced-cars/cheap-lidar-the-key-to-making-selfdriving-cars-affordable

and the providers of technologies for spatial indexing that there will be high-speed connectivity (e.g., 5G) to both centralized and highly distributed servers, with continuous power and optimal conditions to meet the off-device computing challenges.

Besides cloud-based enablers for AR Cloud visual indexing and positioning in real time, there will need to be edge computing resources. The edge computing may be a function of the user's relationships with edge computing providers, the tasks/requirements for the use case, and other factors. We generally assume that edge computing has continuous power supply, and, in either the case of the cloud or edge, the user's device must have uninterrupted high speed and high bandwidth connectivity.

High Performance, Low Cost Mobile AR devices

AR devices come in many shapes and sizes, and have basic connectivity (antennae, etc). To be suitable for AR Cloud spatial indexing and continuous pose tracking/updating, they must also include many of the following key enablers:

- GPS
- Camera with some minimum requirements
- Compass
- Inertial Measurement Unit (IMU) containing gyroscope, accelerometer, and magnetometer
- Barometer
- WiFi
- BTLE
- Radar
- Infrared
- LIDAR
- Cellular 4G/5G

The choices of which sensors to include in a mobile AR display device with AR Cloud capability, and how to use sensor fusion to reduce power consumption, will depend on the use cases and further user requirements.
Some sensors mentioned above are also required for reality capture (see discussion and comparison of sensors above).

High Bandwidth, Low Latency Networks

Using AR Cloud for localization and pose tracking requires connectivity. The precise type of connection (LTE, 5G, WiFi, specialized edge solutions e.g. NB-IoT, etc.) depends on many factors and, provided it meets the speed and bandwidth requirements, is not highly specific.

The minimum speed depends on the use case, the user's device, and the resources needed from the edge or cloud. Likewise, the minimum latency is not a predefined value, since it will depend on many factors. However, latency is likely to be a less forgiving requirement than bandwidth for localization and pose tracking.

The network connectivity may be continuous or intermittent, provided that the AR device has the resources to store and perform the computational tasks on the data it has received from the AR Cloud.

Current Status

At the time of this report, the development of tools and technologies for devices and services to query or retrieve captured information about the physical world, and to position a user in 6 dimensions, is immature and fractured.

The tools are typically provided as part of a larger technology silo that includes reality capture and AR experience display.

Performance of the processes needed for continuously updating the device regarding the registries and its pose in the physical world is also, not sufficiently powerful on mainstream AR display devices (smart phones).

Also, there is no consensus on the formats or syntax and protocols for indexing the physical world and the assets available to the user.

Finally, there is not, to the knowledge of those preparing this report, any proposal or technology for switching between registries or services that AR Cloud persistence continuously needs (for AR Cloud persistence).

Challenges and Opportunities

Similar to the reality capture technologies at the time of preparing this report, the challenges facing spatial indexing and mapping, and geopose include the low levels of interoperability and a low ability to scale given current technologies.

Further, for spatial indexing and mapping, and geopose processes to be delivered in real time by AR Cloud-connected devices requires the key enablers to advance significantly. In particular, many use cases will not be possible until there are:

- Higher network coverage (so that the user can continuously access network-based services)
- Highly persistent and high bandwidth connectivity services, and
- High performance and low power consuming computational resources

Privacy, Security and Data Rights

Geospatial Define the landscape: topical areas
Players: Standards bodies, where they plug into the landscape
Gaps: Especially gaps the OARC might address

OVERVIEW

As our relationship with the Internet and online services matures, the general public understanding of privacy and security issues is slowly improving. Policy makers and advocacy groups are neck deep in discussions regarding data rights and mechanisms for protecting user data. Proposals are being offered and debated about how to provide end-users of online services with the ability to create and manage their preferences, and to help them control the exposure and distribution of their personal data.

AR solutions necessarily involve a suite of technical elements working in concert in order to function properly. The AR solution stack may include the end user's device (smartphone, tablet, smartglasses or headset); the Operating System (OS) and application software; wireless networking functionality; local and remote computing resources; cloud-based data storage, transfer and analytics; AI-driven interactions, reporting and admin applications; integration with Enterprise IT, mobile device management resources, geospatial data sets, and more. Systems engaging AR Cloud resources further increase the number and types of trust boundaries that must be protected.

AR solutions are tightly linked to the environments in which they operate, and sense, process, store, and possibly expose a large range of important information related to personal and business facilities, locations, resources, communication, and activities related to planning, operations/production, maintenance, and more. To be effective, AR systems nearly constantly gather data while in use and even in standby mode. This data can include detailed spatial maps of user surroundings and captured audio, video, locational, and positional data. Some of this data can be accessed remotely without the user even being aware it is happening.

So modern and emerging AR solutions operating in conjunction with AR Cloud resources represent whole new dimensions of concerns related to privacy, security and data rights. Who gets to, or should, experience what AR content, when, where, under what conditions, and how should that content be presented? This section of the report will examine some of the most important aspects of using AR solutions related to privacy, security and data rights. It can only be an incomplete review because the dynamics of the field are changing so rapidly.

- How it fits/Scope

- Key enablers

Current status

Existing manifestos, declarations etc.
"Privacy manifesto for AR Cloud Solutions" released on 18th of October 2018 by OARC when we launched as an organization in AWE EU in Munich
https://medium.com/openarcloud/privacy-manifesto-for-ar-cloud-solutions-9507543f50b6
Ongoing work that has relevance:

- Direction of evolution

- Companies currently doing work in this area

Privacy

Personal privacy is a topic that has always been important. People generally want to control what information about them is exposed to others specifically, and the public generally. Businesses certainly want to restrict access to data about their financial status, human and corporate resources, operational activities, customers, intellectual property. As more of our lives involve digital tools and communication, we each build up large bodies of information about ourselves. Very few of us understand the true extent of Personally Identifiable Information (PII) being collected through the use of our digital tools, by whom, using what techniques, and where that data is being stored and communicated. To what ends are others using our personal data for aims which may or may not be in our best interest? Breaches of our personal and business data can cause serious impact to our productivity, income, efficiency, reputation, safety, security and stability.
As users begin to tap into the potential of the AR Cloud, it is important that solution providers offer layered levels of control and protection to users regarding how their PII is handled depending upon its relative importance. Where payment for services and eCommerce is involved, the situation becomes even more sensitive. There must

also be a method for de-identifying information so it cannot be associated with an individual person, yet can still be mined to yield important insights and valuable data which can be exploited for the broader interests.

The Challenge of Chasing Revenue

One approach for providing users with a greater degree of data control involves offering paid, premium services in addition to free functionality. Paid services mean there is more incentive for the provider to work in the best interest of the individual. Vendors of free services, however, are incentivized to engineer other forms of revenue generation. The differences between paid and free services will probably have a big effect on both Privacy and Security.

Personal Privacy vs. Public Good

Privacy and Security must have a symbiotic relationship in order for either one to be effective. One example is that of the San Bernardino terrorist attack. The terrorists owned iPhones. The government asked Apple to unlock the iPhones in order for justice to be served. Apple said that they couldn't, and if they could they wouldn't. So, the government devised a workaround. That clearly demonstrates the perpetual dance between these two needs. If anonymity is 100% guaranteed, as it currently is with cryptocurrency, then it becomes a haven for criminals and organized crime. If governments mandate back doors into cryptography technology, as they have in the past, then it defeats the point of the cryptography. In order for either Privacy or Security to work, the two need to find a balance.

Rise of Privacy Manifestos and Bill of Rights

Due to the "long-arm juris" nature of regulations such as the GDPR, its best to visualize data protection impacts (positive or negative) on spatial computing and cloud by the following dimensions:

1. how existing laws help protect data rights for AR;
2. enforcement;
3. any existing reported breaches, all in context of users looking for the right to rectify disputes of digital property, right to be forgotten, consent, enforcement/civil action.

82

SECURITY

Data security related to the AR Cloud means adopting or creating standards and tools to protect data from intentional or accidental disclosure or modification. This includes protection across the entire life-cycle of usable data, from creation and collection, storageto storage, transmission, management, retrieval and use. There are many security techniques which can be used to limit access and editing of data by unauthorized or malicious actors. There is still a lot of work to be done in order to understand which of these techniques will be most applicable and practical for AR Cloud solutions.

Generally, ensuring data security in digital systems involves identifying the sources and potential magnitude of security risks, developing strategies for risk mitigation,

Global Spatial Cloud Regulation Map Sources:
https://unctad.org/en/Pages/DTL/STI_and_ICTs/ICT4D-Legislation/eCom-Data-Protection-Laws.aspx
https://www.dlapiperdataprotection.com/
61 Data Breach Survey 2019 (GDPR specific) https://www.dlapiper.com/en/uk/insights/publications/2019/01/gdpr-data-breach-survey/

and striving to reduce the effect that implementing risk mitigation has on AR experience delivery. Certainly, if there is no outside data required by an AR end-user, and no data is collected, stored or transmitted by that user during the course of their experiences, the security risks are small. Obviously, the value of those experiences would be greatly diminished without access to the information and functionality available, otherwise. So solution providers must strike up a balance between risk mitigation and the quality of the experience delivered to the user.

Managing Risks

Security Programs for AR Cloud Solutions

Cyber Security Equivalence Concept

NIST Framework for Improving Critical Infrastructure Cybersecurity
OWASP Mobile Security Project
IEEE Cyber Security Initiative
ISO/IEC Information Security Management Systems Standards
Industrial Internet Security Framework
AREA Security Framework and Test Protocol

Risk Assessments

Identification of Security Requirements

Document Use Cases

Document Design Architecture

Threat Assessment for AR Cloud Solutions

Evaluation of Security Design

Security Protection Levels

Attack Categories for AR Cloud Solutions

Device and Application Protection

Penetration Testing

Data Rights

The topic of data rights is receiving much more attention in the past few months as people begin to realize that creating a world full of extraordinary AR experiences is one thing, however, determining and coordinating who can experience which content, when, where and under what circumstances is quite another. Data rights related to the AR Cloud involve designating who has the legal authority to create, access and use data associated with AR solutions and experiences. In addition to data actually appearing as content in an AR experience, rights must also be understood and respected regarding data used to *generate* the experience, for example, geospatial information, property ownership records, user authentication information and more. Some data is protected from disclosure, use and distribution as Intellectual Property protected under national and international trademark and patent laws. Data Rights focus on the license to use, release, or disclose technical or personal data or computer software to persons who are not otherwise authorized. Data rights are asserted by one party and may be administered, protected and enforced by others. In any realm where data rights are involved, there must also be a means for dispute resolution when conflicts and differences of opinions are involved. Data rights related to the AR Cloud not only involve working out the fundamental understanding of who should have rights to what data, but also in creating the tools for documenting, managing and promulgating those rights, often, in near-real-time. [We might define in common sense terms what people want, even if it doesn't seem they are compatible, we would get closer to the direction we want to go. If we don't define it well, we will have AR graffiti everywhere, it's going to be awful and people won't use it.]

Types of Data Rights

Data Rights Management

Dispute Resolution

Edge Computing and IoT for AR Cloud

Introduction

AR Cloud experiences are inherently mobile and highly location-centric. AR requires significant computational resources, has tight bounds on latency, and can require significant bandwidth. AR Cloud functions such as *reality mapping*, *localization*, and *scene understanding* compound these requirements and introduce a significant data storage dimension. This mix of complex and sometimes competing requirements drives a need for new mobile communications and edge computing.

Scope

Consistent low latency and high bandwidth are key infrastructure requirements for mobile AR experiences and AR Cloud technology.

AR Bandwidth and Latency Requirements

In VR systems, lag up to 20ms can go unnoticed. Over 20ms, humans perceive the misalignment through muscle motion and the vestibular system, resulting in motion sickness. This delay between an action (e.g., head movement) and the corresponding reaction (e.g., display update) is known as the motion-to-photon (MTP) latency. In AR systems, virtual content overlays onto the real world and the threshold drops to 5ms[83].. Interacting with the virtual content in real-time may fall within the scope of Tactile Internet, which also requires latency of less than a few milliseconds.

AR experiences can push bandwidth requirements into the hundreds of Mbps and potentially over 1 Gbps[84].

Additional Requirements

Mobile AR and AR Cloud platforms must support consistent service experiences for highly mobile devices in highly dynamic environments, and maintain scalability, reliability, and security. Offloading compute, memory, and storage from mobile devices can significantly improve size, performance, and battery life. Additionally, mobile AR application developers must be able to access services and resources securely and in a unified manner.

---

[83] DAQRI, MOTION TO PHOTON LATENCY IN MOBILE AR AND VR, August 20, 2018, https://medium.com/@DAQRI/motion-to-photon-latency-in-mobile-ar-and-vr-99f82c480926
.

[84] Chatzopoulos, Dimitris, et al., Mobile Augmented Reality Survey: From Where We Are to Where We Go, Ieee Access, vol. 5, 2017, pp. 6917–6950.

Key Enablers

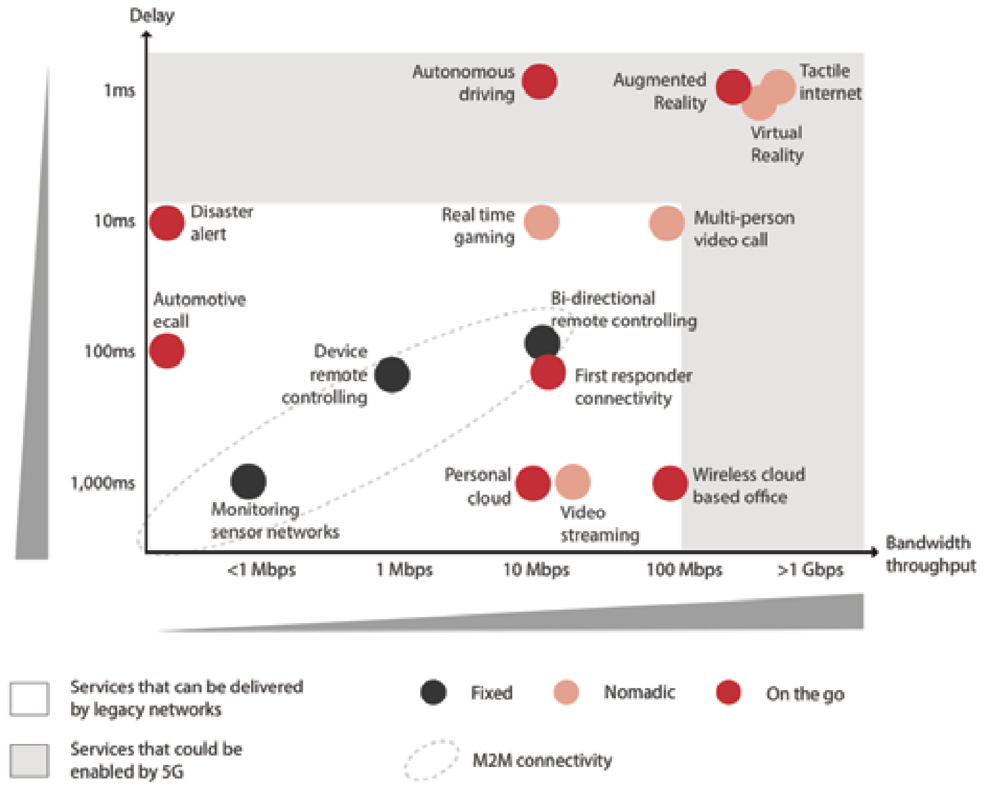
Network Function Virtualization (NFV) and Software Defined Networks (SDN)

NFV involves virtualizing network functions on general purpose (commodity) hardware. This reduces cost, increases efficiency, allows for resources to scale up or down as needed, improves time-to-market, and reduces vendor lock-in. Network functions typically have very strict performance requirements and are often associated with real-time or near real-time services. As such, virtualization must be designed in a manner that guarantees performance, and avoids noisy neighbor problems[85]. Network functions must be designed for scalability and resilience within the virtualized environment.

SDN splits the control plane and data plane in a network to enable logically centralized control of resources. It allows for control, programming, and management of the network resources and enables dynamic, scalable network services. SDN improves time-to-market for network services and allows custom service-aware networking. SDN also allows greater network resource efficiency and on-demand scale up/down.

[85] Wikipedia, "Noisy neighbour" Wikipedia, 18 May 2019 https://en.wikipedia.org/wiki/Cloud_computing_issues#Performance_interference_and_noisy_neighbors

5G Communications



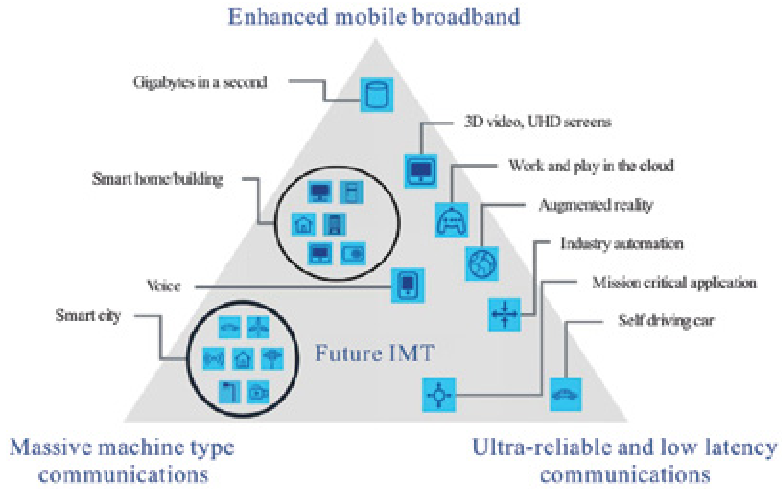
Bandwidth and latency requirements of potential 5G use cases. Source: GSMA Intelligence[86]

As shown in the above figure, 5G enables use cases such as AR and Tactile Internet. The IMT-2020[87] work within the ITU sets targets for 5G. This includes a 100x improvement in bandwidth (to 10 Gbps) and a 30-50x improvement in latency (to 1ms)[88].Dense small cell deployments, the new millimeter wave (mmWave) band, massive multiple-input, multiple-output (M-MIMO), and beamforming have all enabled key advanced in 5G[89]. Indoor/outdoor dense small cell deployments allow macrocells to be offloaded, improve spectrum reuse, increase network capacity and reduce latency. However, this can also drive higher handoff rates and inter-cell interference. New frequency bands are required to address the 5G demands. The mmWave band uses frequencies in the 30 to 300GHz range. mmWave has very different propagation characteristics limiting cell radius to <100m and creating

[86] GSMA Intelligence, Understanding 5G: Perspectives on Future Technological Advancements in Mobile, 2014, pp. 1–26.
[87] ITUR WP5D, Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond, 2015
[88] NGMN.Alliance, 5G White Paper: Next Generation Mobile Networks, 2015, pp. 1–125.
[89] Al-Falahy, Naser, and Omar Y. Alani., Technologies for 5G Networks: Challenges and Opportunities, IT Professional, vol. 19, no. 1, 2017, pp. 12–20

problems with objects (including the human body) that block line of sight. M-MIMO involves deploying large numbers of antennas at base stations to improve signal strength and throughput, while beamforming involves concentrating power into a beam with limited width but a large gain. There are many challenges with M-MIMO including physical size, contamination from nearby cells, and accurate channel estimation.

IMT-2020 defines three major service categories: enhanced mobile broadband, massive machine type communications, and ultra-reliable and low latency communications (URLLC). Given the bandwidth and latency requirements for AR, it falls in between enhanced mobile broadband and URLLC. While some 5G networks have already launched, standards around low latency are still in development. Specifically, the URLLC service category in 5G depends on the 3GPP New Radio (NR) air interface, which is currently targeting for commercialization in 2020[90].



Usage scenarios of IMT for 2020 and beyond. Source: M. Series [91]

Besides significant bandwidth and latency improvements, 5G also provides other benefits to AR Cloud. For example, the mmWave band of 5G has the potential to assist in mapping and localization, augmenting vision-based mechanisms, as discussed in non-visual localization for WG-1. 5G networks introduce device-to-device (D2D) communications, which allow end terminals to communicate directly without traversing the mobile network. D2D provides performance benefits, but it also creates new opportunities. Deployment of fixed terminals with known location can

[90] Ji, Hyoungju, et al., Ultra-Reliable and Low-Latency Communications in 5G Downlink: Physical Layer Aspects, IEEE Wireless Communications, vol. 25, no. 3, June 2018, pp. 124–30
[91] Series, M., IMT Vision–Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond, 2015, pp. 2083–90.

assist in indoor/outdoor localization via D2D. Establishing ad hoc and peer-to-peer local networks allow new types of services, provide network connectivity in the event of infrastructure failures or loss of coverage, or to aggregate many smaller sensors in an IoT setting with efficiency.

Edge Cloud Players

The most well-known cloud players are the tier-1 public clouds such as Amazon Web Services (AWS)[92], Microsoft Azure[93], and Google Cloud[94]. These clouds have some level of geographic distribution to allow users to keep data within a region and to improve regional latency and bandwidth. However, they are generally regarded as "centralized" cloud. There are some specific use-cases in which public clouds can have greater distributiontoday. These include Internet-of-Things (IoT), Content Delivery Networks (CDNs), and hybrid cloud arrangements.

IoT extends the Internet to physical devices other than traditional computers or mobile phones. IoT devices span many consumer applications including smart homes, wearable computing, and connected vehicles, and include devices in commercial, industrial, and infrastructure applications[95]. Cisco estimates that about 50 billion IoT devices will be added to the Internet by 2020[96]. These devices often have limited resources (compute, memory, storage) and limited communications capabilities. AWS IoT Greengrass[97], Azure IoT Edge[98] and Google Cloud IoT[99] bring cloud capabilities to local devices, and allow for localized communications while potentially disconnected from the public cloud.

Content Delivery Networks (CDNs) are more distributed than public clouds and primarily operate as web caches that accelerate delivery of content to end users. Pure CDN companies such as Akamai[100] and Limelight Networks[101] provide CDNs, but also do public clouds, such as AWS CloudFront[102], Azure CDN[103], and Google Cloud

[92] https://aws.amazon.com/ .
[93] https://azure.microsoft.com/en-us/
[94] https://cloud.google.com/
[95] Wikipedia, "Internet of Things" Wikipedia, May 4, 2019, https://en.wikipedia.org/w/index.php?title=Internet_of_things&oldid=895493198.
[96] Evans, Dave, How the Next Evolution of the Internet Is Changing Everything, 2011, p. 11.
[97] https://aws.amazon.com/greengrass/
[98].https://azure.microsoft.com/en-us/services/iot-edge/.
[99] https://cloud.google.com/solutions/iot/
[100] https://www.akamai.com/
[101] https://www.limelight.com/
[102] https://aws.amazon.com/cloudfront/
[103] https://azure.microsoft.com/en-ca/services/cdn/

CDN[104]. More recently, the rise of streaming video has led telecom operators to develop their own CDNs. Telcos have the additional benefit of owning the networks that the video content traverses. CDNs traditionally dealt with more static content, but today CDNs offer some edge application functionality such as Akamai Cloudlet Applications[105] and AWS Lambda@Edge[106].

Large enterprises often run on-premises clouds such as VMware[107] or OpenStack[108] to reduce costs and meet specific security and regulatory requirements. Hybrid cloud arrangements that bridge on-premises clouds with public clouds are also becoming commonplace, with integration initially focusing on private network arrangements such as AWS Direct Connect[109], Azure ExpressRoute[110], and Google Cloud Dedicated Interconnect[111]. Some public clouds are offering on-premises solutions to support hybrid environments. For example, Microsoft Azure Stack [112]and Google Kubernetes Engine On-Prem[113] allow some cloud services to run on-premises and he standard cloud portals and APIs to manage them. These hybrid arrangements could provide a path for public clouds to move much closer to the edge. Public clouds offer many services to application developers today. A key challenge will be determining which services will most benefit from an edge presence, and how they will eoperate in a highly distributed model.

Telco clouds today are very much a work-in-progress, focusing largely on 5G and Mobile Edge Cloud (MEC). Telco clouds are designed to be highly distributed and close to the mobile users, making them an excellent fit for AR Cloud.

Edge Computing

Edge computing encompasses many paradigms including mobile cloud computing (MCC), fog computing, cloudlets[114], and MEC. The different edge computing paradigms overlap in terms of goals and technology characteristics.

104 https://cloud.google.com/cdn/
105 https://www.akamai.com/us/en/products/performance/cloudlets/
106 https://aws.amazon.com/lambda/edge/
107 https://www.vmware.com/
108 https://www.openstack.org/
109 https://aws.amazon.com/directconnect/
110 https://azure.microsoft.com/en-us/services/expressroute/
111 https://cloud.google.com/interconnect/docs/concepts/dedicated-overview
112 https://azure.microsoft.com/en-us/overview/azure-stack/
113 https://cloud.google.com/gke-on-prem/
114 Wang, Shuo, et al., A Survey on Mobile Edge Networks: Convergence of Computing, Caching and Communications. IEEE Access, vol. 5, 2017, pp. 6757–6779.

MCC considers specific aspects of mobile devices for offloading, including bandwidth, connectivity, energy, location, and mobility. However, the long distance between edge devices and clouds in MCC makes it unsuitable for AR applications.

The analogy that "fog is closer to people than clouds" is the basis of *fog computing*. It focuses on the IoT and offloads computing to edge devices such as routers. The OpenFog Consortium[115], founded by ARM, Cisco, Dell, Intel, Microsoft, and Princeton University, backs fog computing. Fog computing includes collaboration between end user clients or near-user edge devices. This also enables operation in cloud-disconnected scenarios, which can be common for IoT. Fog computing is a more generic architecture when compared with MEC, allows for a hierarchy of fog layers, and is not mobile operator centric.

Cloudlets originated in academia as a representation of a "data center in a box"[116], aimed at moving existing clouds closer to the edge. A cloudlet typically has a specific role and allows offload from mobile devices. However, it only contains soft state which makes it easier to manage, but assumes robust connectivity back to the primary cloud[117].

5G and MEC

MEC is being standardized by ETSI[118] and 3GPP[119], and is part of the 5G Public Private Partnership (PPP)[120]. MEC and the alignment with 5G is especially relevant to the AR Cloud. It involves collocating cloud servers with base stations, transmission nodes, network aggregation points, or with core network functions, for highly efficient offloading. MEC provides low latency, proximity, high bandwidth, real-time radio network information and location awareness[121].

The core facets of ETSI MEC are application enablement, API principles, service related APIs, and management and orchestration APIs[122]. Application enablement involves registration and service discovery, service authentication and authorization, and services communication. API principles ensure a consistent set of developer-friendly APIs. Service related APIs expose network and context information, and

---

115 https://www.openfogconsortium.org/.

116 Wang, Shuo, et al., A Survey on Mobile Edge Networks: Convergence of Computing, Caching and Communications. IEEE Access, vol. 5, 2017, pp. 6757–6779.

117 Elijah Home. http://elijah.cs.cmu.edu/

118 https://www.etsi.org/technologies/multi-access-edge-computing .

119 https://portal.3gpp.org/#/

120 https://5g-ppp.eu/

121 Wang, Shuo, et al., A Survey on Mobile Edge Networks: Convergence of Computing, Caching and Communications. IEEE Access, vol. 5, 2017, pp. 6757–6779.

122 Filippou, Miltiadis. ETSI MEC: Where Are We Now. 2018, p. 35.

management and orchestration APIs allow applications to dynamically invoke in the appropriate location.

The synergies between 5G and MEC are an important aspect. Both 5G and MEC specifications contain the same Service Based Architecture (SBA) concepts, which involve alignment of system operations with NFV and SDN. MEC and 5G systems can collaborate in traffic routing and policy control. In SBA, there are service producers and consumers, and consumers are authenticated and authorized to use specific services exposed by producers. The framework additionally provides service registration, discovery, and notifications. 5G provides network slicing which allows different services to allocate resources. An application in MEC can associate with one or more network slices. The resources in an MEC instance are tiny compared to centralized cloud, so efficient resource reservation and traffic steering for endpoints that are potentially mobile is especially important. MEC can control traffic steering for a single endpoint or group of endpoints. For stateful applications, when an endpoint moves into a new MEC area, the state must transfer and synchronize in careful coordination with the application logic. MEC is able to subscribe to endpoint movement notifications from the 5G network. This is an example of the wider capabilities exposure framework.

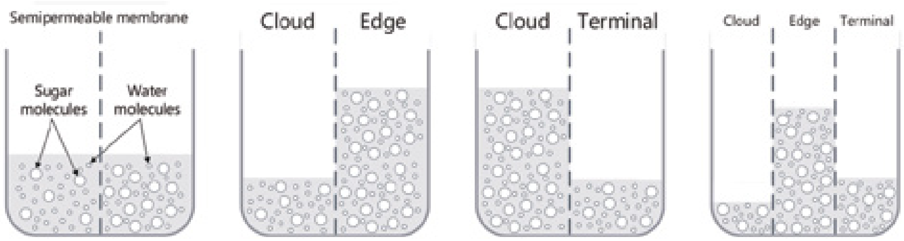
Computation and Data Offload and Locality

MEC allows data and compute to offload from centralized cloud in a highly distributed manner with consideration for the locality of mobile clients.

The proximity of MEC to the mobile clients combined with 5G communications also allows for data and compute to efficiently offload from mobile clients to MEC. This is important, as mobile devices have CPU/GPU, memory, and storage limits. Furthermore, aggressive use of local compute resources can significantly impact battery life.

Offloading computation and data to MEC is highly beneficial to AR Cloud. Reality mapping (including scene understanding) and localization have an inherent location affinity. Potentially, MEC may also handle virtual content and rendering.

Osmotic Computing



Osmotic computing and sample scenarios. Source: Xiuquan Qiao[123]

Osmotic computing[124] is a new paradigm that automatically distributes microservices in an adaptive and scalable manner. We expect the AR Cloud to follow a model in which microservices are distributed between centralized cloud, edge cloud, and the terminal, as in the above figure.

Programming at the Edge

A key challenge for edge cloud is creating an abstraction layer that masks provider and implementation differences and offers a consistent interface to application developers.

MobileEdgeX[125] is a Deutsche Telekom-founded company that uses open source software to aggregate highly distributed edge infrastructure and deliver normalized, abstracted interfaces to application developers. MobileEdgeX is live in production in Germany, allows application backends to be deployed close to users based on verified location and identity, provides AR and MR performance support, and enables video and image processing.

[123] Qiao, Xiuquan, et al., Web AR: A Promising Future for Mobile Augmented Reality–State of the Art, Challenges, and Insights, Proceedings of the IEEE, 2019.
[124] Villari, Massimo, et al., Osmotic Computing: A New Paradigm for Edge/Cloud Integration, IEEE Cloud Computing, vol. 3, no. 6, 2016, pp. 76–83.
[125] https://mobiledgex.com/

Akraino Edge Stack[126], a Linux Foundation project started by AT&T and Intel, is an open source software stack for edge computing. It seeks to develop edge middleware, SDKs, edge APIs, frameworks for interoperating with other clouds and an application ecosystem. The current members also include Arm, Dell EMC, Ericsson, Huawei, inwinSTACK, Juniper Networks, Qualcomm, Nokia, Radisys, Red Hat, Seagate, and Wind River.

OpenEdgeComputing (OEC)[127] focuses on applications, developer APIs, and seeks to expose standardized and open edge resources. It works alongside ETSI MEC and its membership includes Carnegie Mellon University, Intel, Microsoft, Crown Castle, Vodafone, Deutsche Telekom, NTT, and Nokia.

Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), Containers-as-a-Service (CaaS), and Functions-as-a-Service (FaaS)

IaaS is a basic low-level virtualization service which provides virtual machines (VMs) to users. The user is ultimately responsible for the configuration and management of the VMs, their operating systems (OSs), the application runtime, the application, and associated components such as databases. While VMs can instantiate or de-instantiate as needed, this is a process that can take several minutes, and is not generally well-suited to highly dynamic environments. IaaS technologies include OpenStack and VMware vSphere, and public cloud offerings include AWS EC2[128], Azure VMs[129], and Google Compute Engine[130]. IaaS may provide the internal foundation for other service offerings in an edge cloud setting, but demand for external IaaS offerings may be limited.

Containers provide virtualization of an OS and align well with the concept of designing applications as a collection of microservices. Containers can instantiate or de-instantiate rather quickly, making them ideal for highly dynamic environments. The rise of Docker largely popularized the concept of containerization.r[131]. Kubernetes has more recently gained traction as a container orchestration platform[132]. CaaS provides a hosted environment that manages container orchestration and the associated underlying infrastructure, potentially IaaS. In a CaaS environment, users are still responsible for managing container images which include the application, the application runtime, and system tools and libraries (as opposed to a full OS).

126 https://wiki.akraino.org/display/AK/Akraino+Edge+Stack
127 https://www.openedgecomputing.org/
128 https://aws.amazon.com/ec2/
129 https://azure.microsoft.com/en-us/services/virtual-machines/
130 https://cloud.google.com/compute/
131 https://www.docker.com/
132 https://kubernetes.io/

Examples of CaaS include Amazon Elastic Container Service[133], Azure Kubernetes Service[134] and Google Kubernetes Engine[135]. CaaS is popular with application developers using dynamic microservices who require some lower-level control. Edge cloud environments are likely to offer CaaS, and would be well-suited to address components of an AR Cloud architecture.

PaaS provides additional abstraction over IaaS and CaaS and allows application developers to avoid configuration and management of VMs, OSs, containers, application runtimes, or associated components such as databases. Some PaaS examples include Heroku[136], Google App Engine[137], and AWS Elastic Beanstalk[138]. The lightweight nature of PaaS makes it ideal as an edge cloud offering, and a viable option for AR Cloud development.

Finally, FaaS is a relatively new entrant, associated with the concept of serverless computing. FaaS is like PaaS, but includes some higher level abstractions and auto scales on a per-request basis. FaaS is good for event-based or streaming applications commonly found in IoT settings, but is not great with long running persistent processes, network processing, or databases. FaaS was born through the launch of AWS Lambda[139] in 2014. Microsoft and Google have since launched Azure Functions[140] and Google Cloud Functions[141] respectively, and open source options such as OpenFaaS[142] and OpenWhisk[143] have also emerged. FaaS offerings are especially relevant in highly dynamic edge cloud settings for the appropriate use-cases.

Challenges and Opportunities

5G standards are still evolving and 5G networks are in the very early stages of deployment. Additionally, 5G networks involve multiple radio technologies with different characteristics. When considering multiple telecom operators with different phased 5G rollouts for different areas, consistent service experiences may take several years. Service characteristics within dense urban areas may diverge

133 https://aws.amazon.com/ecs/
134 https://azure.microsoft.com/en-us/services/kubernetes-service/
135 https://cloud.google.com/kubernetes-engine/
136 https://www.heroku.com/
137 https://cloud.google.com/appengine/
138 https://aws.amazon.com/elasticbeanstalk/
139 https://aws.amazon.com/lambda/
140 https://azure.microsoft.com/en-us/services/functions/
141 https://cloud.google.com/functions/
142 https://www.openfaas.com/
143 https://openwhisk.apache.org/

from those outside of these areas at a greater level than seen with previous mobile networks.

Edge cloud is also a very new and rapidly evolving technology. There are multiple types of cloud players including centralized public cloud, CDNs, IoT clouds, highly distributed telco clouds such as MEC, and on-premises clouds. Different cloud players have different characteristics and are ultimately suited to different types of applications (or even sub-components of applications). Centralized public clouds provide the richest ecosystems today with extensive portfolios of cloud services. However, the lack of edge distribution presents major challenges for the highly localized and low-latency applications associated with AR Cloud. Telco edge clouds such as MEC address the locality and low-latency, but they are still in the early stages of infancy, far behind centralized clouds in terms of rich cloud services, and lacking unified APIs that allow for provider-agnostic and infrastructure agnostic application development.

Compliance, Regulations, and Legal

Introduction

The state of the AR Cloud regarding compliance regulations and legal issues is, at best, incomplete. Current policy has largely been based on decades of commercial and consumer regulations that vary widely due to laws and protections established by local, regional national and international jurisdictions. We will need to create or adapt current policies, regulations, or legal frameworks to provide clear guidance on issues surrounding the creation and consumption of this new, digital reality. This includes reality capture, content creation, virtual goods, licensing, intellectual property, and most importantly, the protection of user privacy, rights and opportunity.

The legal landscape governing reality capture, one of the fundamental prerequisites for AR Cloud technology, is a patchwork of overlapping and inconsistent laws, rules, and regulations. Few, of these rules were written with anything like the Open AR Cloud in mind. The technology implicates multiple levels of privacy, intellectual property, and related concerns. These topics overlap with the topic explored by other OARC working groups, which is to be expected. The OARC working group on Compliance, Regulation and Legal's main contribution lies in its focus on aligning these various topics with existing legal frameworks. This document serves as an initial high-level summary of the biggest challenges rather than attempting a comprehensive overview or plan of action. We must work further in the future to get a more complete picture and deeper insights.

Main Goals

This group's fundamental purpose is to align what the OARC does with existing frameworks of law and regulation. We cannot accomplish such a broad goal immediately, so we undertake it in successive increments. The first pass should identify the primary technological tasks necessary to the existence of an OARC, the primary jurisdictions in which this will take place, and the laws we must comply with in order to achieve these tasks. The second layer of analysis will describe a more complete picture of OARC technology, interactivity, content, and the broader range of rules we must observe. Third, this group should identify legal doctrines and texts that we should adopt, modify, or eliminate to be in alignment with OARC's vision for the Open AR Cloud ecosystem.

State of Reality Capture

“Reality capture”, as it relates to the AR Cloud, are collections of types of data anchored geospatially to a real world frame of reference and stored in databases. Reality is mapped and captured in different ways by gathering observations or sensor data from the physical world and associating that data with its real world anchors.

At its most fundamental and abstracted level, this process is indistinguishable from 2D mapping technology that has existed for millennia. What makes OARC-level reality mapping unique is its 3-dimensional nature, millimeter-level precision, information density, and constant (ideally, real-time) updating to reflect changes in real-world conditions. These characteristics are necessary to achieve a robust framework for overlaying the physical world with valuable, usable AR content.

Such a level of reality capture, however, has not quite become a reality yet at least not at a global scale at such a level of detail. However, the technology for real-time mapping of immediate surroundings exists and is maturing rapidly. Efforts are underway to collect this data in centralized repositories. OARC warns against such efforts as it constitutes an unhealthy concentration of data wealth with inherent risks abuse by those in possession of such datasets. This could easily be to the detriment to the liberty and freedom of all people. OARC instead proposes distributed decentralized approaches to managing the data from reality capture. But a robust framework for this data does not yet exist.

Partners and Related Efforts

Other than the OARC itself, many companies are pushing the bounds of reality capture capabilities forward. Magic Leap, Microsoft, Google, DAQRI, Facebook, and several other hardware manufacturers produce AR headsets with impressive “room capture” capability. Software companies such as 6D.ai are likewise pushing these boundaries.

Google is probably the leader in collecting up-to-date, robust data about physical locations. Niantic is the leader in location-based AR experiences. Experts at these companies would make for productive partners in expanding our understanding of the challenges ahead.

Main Challenges

The primary challenge facing this group is the sheer enormity of the task. We can tackle that obstacle by following the iterative process outlined above—by starting with the big picture and proceeding one step at a time from there.

The second challenge will be to develop a detailed understanding of what OARC needs to succeed. Third will be to identify the relevant legal rules. To accomplish this, we must first limit our scope to particular jurisdictions, since the laws of various countries will differ and often conflict with each other.

Call to Action

This is a big but important task, and it cannot succeed without the participation of those who truly know what they are talking about in these fields. Coming up with a platform that does not reflect the reality of the technology or of the law would be a worse outcome than having no platform at all. That is why efforts such as those going in within OARC is important. But OARC also hopes that a much wider range of stakeholders take part in discussion around legal, regulatory and compliance issues related to AR Cloud technology in particular and spatial computing in general.

User Experience, Accessibility, and Safety

Augmented Reality and the Need for Quality UX

New technologies are commonly impelled by the demands of the advancement itself and rarely by the needs of the intended user. Augmented Reality is no exception. As a result, the types of technology available often drive experiences rather than what humanity will from benefit most. With great foresight, OARC chooses not to repeat history, but take advantage of the wide adoption of user experience in modern-day product design. By including experience design, accessibility and safety within our foundation principles, OARC intends to weave solid user experience principles throughout all working groups and create an AR Cloud that is open, accessible and usable for all humankind.

Evans and Koepfler[144] defined AR as, "a human experience with the world that is enhanced by technology." They say, "AR should be less about technology doing things to [and for] people and more about people engaging with the world around them, and having that world enhanced by technology where and when appropriate." In this section, we will show examples of companies, researchers, and enthusiasts enhancing the world in appropriate, and maybe some not so appropriate ways. We will focus discussions about using AR's powers for good, but we will also point to darker possibilities to raise awareness, so we as a collective community can lead the charge to avoid as many unintended consequences as possible.

And likewise, Angie Li and Therese Fessenden, from the Nielsen Norman Group[145] see AR as a tool that can improve user experiences. In their article, *Augmented Reality: What Does It Mean for UX?*, they show that a well-designed interface that displays useful information in context can decrease the interaction time to complete a task, decrease the cognitive load and minimize attention switching by combining multiple information sources into a single experience. This WG will look at a breadth of practitioner's work such as this to synthesize a platform of best practices. It is a starting point, based on what we know now. Our work will be to evolve the best practices platform alongside the exponential changes we expect to see as the AR Cloud evolves.

---

[144] Evans, Kieran and Koepfler, Jes A., The UX of AR: Toward a Human-Centered Definition of Augmented Reality, http://uxpamagazine.org/the-ux-of-ar/

[145] Li, Angie and Fesseden, Therese, Augmented Reality: What Does It Mean for UX?, September 18, 2016, https://www.nngroup.com/articles/augmented-reality-ux/

Charter of User Experience, Accessibility and Safety Working Group

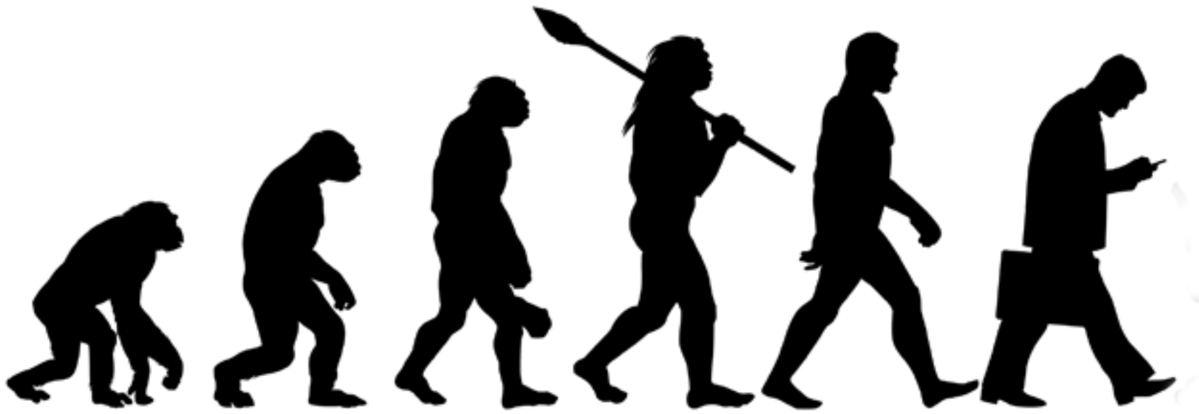
We have an unparalleled opportunity to build a better future for all of humanity. At its core, AR is a tool to augment our world, our lives, our reality. With intentional forethought AR, and the AR Cloud, is a medium that can make the world more accessible, easier to understand and use. This WG is committed to create an AR Cloud that reduces barriers, increases inclusivity and makes the world itself a better user experience for all.

Adoption of Augmented Reality: A Business Case for UX

> "July 2016, Niantic, the parent company of PokemonGO (and of the popular augmented reality game Ingress) reported revenue streams of $10 million per day from PokemonGO alone, proving that augmented reality features can be successful in a mainstream market." [146]

Yet, we haven't seen a massive wave of adoption. Why not?

CEO of Kopin, Dr. John Fan posits that adoption will depend on the evolution of the headset and the value of an experience's function. He talks about the millions of years of evolution it took for the human species to stand upright, to look out into the world, head held high with a favorable vantage point. Looking down on our phones is an evolutionary regression that has taken us out of the real world. Dr. Fan also asserts that AR and wearables need to bring us back in to the real world--that the devices need to be on our head near our eyes and our ears so we may look back out into the world as evolution intended.



[146] Ibid

Dr. John Fan asserts that devices for augmentation need to return us to an upright and natural position.

Fan is also emphatic that people don't want to wear anything on their head and especially not a piece of technology. So how do we overcome this dilemma? The headset must be comfortable, stylish and provide a compelling function. We wear glasses to correct vision, but also select them based on style. Sunglasses, hats, helmets all provide an essential function. Therefore creators of AR headsets must create an incentive for people to put them on, and it must captivate enough for people to keep them on. It must be something that improves their world. As the value provided by the headset increases, the willingness to wear it will also increase.[147]

Dr. Fan is making a case for creating a well thought out user experience. He makes both a cultural case, improving the world, and a business case, increased demand. Paul Boarg, author of *User Experience Revolution* writes:"Your customers define your brand identity. They decide what you are like through their social media updates, through reviews, through all of the different channels they've got – they decide how your brand is perceived… And as a result, producing an outstanding experience is one of the best ways of ensuring a positive brand identity."[148] A good user experience isn't just the 'right' thing to do, it also makes shrewd business sense.

Scope

Augmented Reality is the interaction of superimposed data, graphics, audio and other sensory enhancements over a real-world environment — the world we actually see, the world within which we actually work, the world our citizens navigate every day. This differentiates AR from virtual reality, which places the user in an entirely created, virtual world. The experience of AR is simple, but powerful — it is contextual, visual and even visceral.

When thinking about the relationship between concepts and technologies of an AR Cloud, some important questions arise. Some seem to have answers already, others require discussion and research from this WG. It's important to establish a solid UX process to help drive the WGs direction and vision. The following is therefore best understood as a beginning rather than a result. In line with this view, there are three core pillars to help direct our objectives within this WG:

---

[147] Fan, J., Dr. Fan's Rules for Successful AR and VR – Inspiring Success, Youtube, June 21, 2018, https://www.youtube.com/watch?v=W9tx9nDUOys

[148] Ratcliff, C., Interview: Paul Boag on the "User Experience Revolution", March 28, 2018, https://www.userzoom.com/blog/interview-paul-boag-user-experience-revolution/

Pillar One: Establish Use Cases in Collaboration with other Open AR Cloud Working Groups

The first aim of this WG is to create a clear and well thought out research approach. As you might imagine, there is a lot of experimental and theoretical work happening in the Immersive Design space across a wide spectrum of industries. We believe the most impactful approach is to work in collaboration with the other working groups to identify the primary concerns amongst the groups and establish a set of well-formulated use cases we can prototype and user test. In addition, we will collectively, amongst all working groups, identify, analyze, and practically apply what we have learned in the research and writing of this report to the WG's research.

This approach empowers us to identify audiences and use cases that will help express the best-in-class and most forward thinking design in response to several of AR Cloud's most potentially demanding challenges. Our WG's focus on UX is critical to all aspects of the Open AR Cloud community, so collaboration and development with these other groups is vital to a healthy Open AR Framework. We believe that quality User Experience is tantamount to the development and adoption of safe, accessible and ethical Augmented Reality technologies.

Pillar Two: Prototype and Test

The second pillar is to conduct our own experimentation across a variety of tools, technology and experiential use cases that have a direct impact on people's everyday lives. In doing so, we can identify emergent and dominant UX patterns and take initial steps to validate the effectiveness of these patterns.

Because of the immaturity of the AR ecosystem, a standard set of tools are still not widely available to address all the existing and expected needs for designing, building and prototyping all the use cases of concern. Despite this challenge, we will devise prototyping methods to approximate the experiences and incrementally validate subsets of use cases until which time complete solutions become more available. We have identified several tools for use in tandem in this paper's appendix *WG8-AP.1*. We welcome suggested additions at oarc.wg@gmail.com. As a by-product of our explorations, we hope to foster innovative and unique approaches to prototyping and user testing for the AR Cloud.

To help support a healthy Open AR Cloud community, we have compiled a collection of comprehensive use cases in appendix *WG8-AP.2* and will continue immediately following the publication of this report. These use cases can help the AR community apply a baseline to the current capabilities available to the AR Cloud and other AR platforms. After our working group has selected a new use case for testing, we will pick the best tools available for the research project. We will work hard to reach a population as diverse and inclusive as possible to represent each persona to validate the thesis and findings of each use case. This will help to confirm assumptions as the working groups propose new and novel approaches to the AR Cloud.

The process this WG uses to and prototype and test is constantly evolving as the entirely volunteer organization works towards the collective mission of advancing usable and inclusive AR technology throughout the world.

Pillar Three: Document and Disseminate

Last, it’s important to document, publish and share the findings of our research and to summarize and share research happening around the world relevant to UX. This allows us to take the best-in-class experience design recommendations and expand access to, and hopefully influence, more creators in meaningful ways to address our concerns for user experience, safety and accessibility for the AR Cloud.

We believe it’s important to educate and help others learn from the foundational work the group has been conducting to standardize both a vernacular and working terminology to further help speed and standardize conversations around complete AR Cloud concepts. Our glossary is an adaptation of existing and new terms and a collective knowledge pool of hundreds of individuals.

We believe this broad perspective will help others outside the industry learn why it’s important to understand how experience design directly impacts implementation for the AR Cloud. In addition, our use cases are a sampling used to to help us guide and drive new thought and will enable us to think holistically about how we train a new class of UX design professionals.

Shared Environments for Collaboration

We are a collective of individuals spread around the globe where meeting in-person regularly is impossible. Through AR shared experiences, we can collaborate in real time without travel time. While we use digital technologies today mostly alone, AR Cloud supported environments encourage the meeting and collaboration of multiple people. Objects of interest are naturally accessible to multiple users. Sharing and editing content together, as we know it from modern online collaboration tools (GoogleDocs, Trello etc.), will spread to spatial situations and offer many new opportunities. This Working Group’s goal is to foster this adoption, when applicable, and be a champion on the benefits to the community.

Key Enablers

As an organization, we have identified several key areas that build upon each other and make our work possible. It's important to understand how these areas will shift and change across varying iterations of guidelines, emergent and dominant interaction patterns, and future technology. We have leveraged similar industries to help construct a practical foundation beyond theory. These enablers are meant to lay foundational thought and groundwork for WG8's research and use cases. The following areas will involve crucial evolution to help enable the group and its future objectives:

Developing a Well Designed AR Cloud

It's important that each of the working groups understand an "accessibility-first" mindset when thinking about creating and enabling these types of experiences. The wide availability of AR-capable devices are only being used in small populations. This poses a crucial challenge as our spaces are dynamic and our ability to understand and analyze data for accessibility use is often limited in sampling scale.

As a group we believe that a central hub of common principles and identifying technology capabilities will be crucial to help bridge the gap for an accessible AR Cloud. This will require some large-scale research as well as using multiple technologies such as AI, Computer Vision, Neural Network processing, and capability classifications as a particular device is accessing the AR Cloud layers.

Several new and existing tool modifications we will be necessary for maintaining the integrity of the AR Cloud including state changes such as creating, updating, archiving, deleting, classifying, and handling meta state layers. We need to know about tools that will allow data retrieval, and browsing. The assumption is that these tools are open and allow the user to access layers of data, information architecture, services, etc. available in the AR Cloud. This includes backwards integration into existing interface systems (mobile, and desktops, IOT, HMDs.)

We will need additional tools for enabling cross-platform access and accessibility hooks into the AR Cloud. This also includes thinking about accessible devices would handle visual, aural, haptic and other sense feedback to adjust for a particular user's specific needs.

Last, we will need new tools and approaches to help manage security protocols across private and public shared data. This includes the ability for a user to determine personal and preferential AR Cloud based profiles to limit or classify personal and private data. For more on security please see the *Security and Privacy* section.

Content Guidelines and Interoperability

This group believes an open set of guidelines will allow everyone to unite around standards for publishing, viewing, and modifying content, as well as service architecture for varying AR Experience Layer Themes. These themes will help to classify alike and distinguish varying content in a way similar to traditional information architectures, but on a world-level three-dimensional spatial scale. For more information, please read the *Content and Delivery* section**.**

Scene Descriptors

Similar to HTML tags, ARIA, and page content descriptors, how we choose to describe 3D scenes, content layers, objects and their positions will impact our ability to create accessible worlds and overlays. This is a crucial information architecture to get right, and many factors for usability, accessibility, safety, privacy and security will be at stake. We must bring together a wide range of expertise to address all these concerns effectively.

Public Services Connected

AR will serve as the "visual portal" to data across the public and private sectors, adding huge value to the prospect of data as a true public asset and resource. With the combination of smart infrastructure: embedded sensors, actuators, displays and computational elements, big data and open data, public sector entities at all levels can stitch together the fabric for Smart Cities. As connected and cross-platform solutions, they deliver integrated services and experiences to citizens and allow workers to operate in that kind of environment. Curtin[149] envisioned several use cases for AR in Smart Cities:

- **Public Safety and Emergency Services**
    - AR applications that provide audiovisual guidance for citizens seeking refuge, evacuation routes, or emergency help in a disaster situation.
    - Citizens and businesses can access authorized geo-specific data on crime statistics and other environmental factors just by pointing their mobile devices at a building, down a street, or for an entire community.
- **Public Health, Wellness and Sustainability**
    - Inspectors of all kinds — health, building and public safety, environmental quality, etc. — can instantaneously "see" and interact with all the information related to a facility, an agricultural area, a district.

[149] Curtin, G., Six ways augmented reality can help you see more clearly, Smart Cities Council, https://smartcitiescouncil.com/article/six-ways-augmented-reality-can-help-you-see-more-clearly

    - A host of environmental quality (air, water, ground, etc.) detectors displayed in AR allows environmental officials and citizens to make real-time decisions on movement, activity and official response.
- **Transportation and Urban Mobility**
    - Imagine in this augmented future being able to see and visually "connect" the various transportation systems — from traditional highway, roadway and fixed-rail infrastructure, to modern on-demand and shared mobility services and active transit (i.e,. walking and biking).
- **Culture, Heritage and Tourism**
    - Across the globe, protection of heritage and culture is a high priority. One of the richest uses of AR is to enhance places, such as historic buildings, castles, monuments and heritage sites and battlefields.
    - Museums and culturally significant buildings are perfect candidates for AR information and rich content around artwork, artefacts, publications, etc.
    - Natural resources - like National parks - combined with AR can provide a powerful educational experience while simultaneously encouraging and monitoring the use and preservation of natural resources.

Curtin concludes with the call: "Tie this all together — wayfinding, things to do, art and culture, history and heritage — to create compelling connected AR experiences for cultural tourists and citizens alike." We emphasize that not only is an Open AR Cloud a necessity for such applications but also that this level of interconnectedness and cooperation with the local governments will be necessary to develop an accessible, safe and inclusive user experience to benefit all.

It's important to understand the impact that places and physical environments will have on defining how we interact in these "smart spaces". We will need to adapt to support new ways of working with the AR Cloud. This includes understanding the functionality, aesthetic, usability, organization, and arrangement of the places we live. Understanding the user behavior expected in these spaces is crucial to thinking about the capabilities, and accessibility concerns, that could end up driving how the AR Cloud works in different environments.

Designers United

Environmental design concerns itself with how the arrangement, functionality and aesthetic of the physical surroundings in which people live, work and play contribute to the quality of the human experience. It considers the symbiotic impact that buildings, parks, natural scenery, communities, neighborhoods, transportation systems and other infrastructure have on each other and how they affect, and are affected by, human behaviour. This approach appears to be a helpful starting point and mindset for integrating AR into environments and for designing the spatial web UX.

Thankfully, some urban designers are looking beyond designing merely for safety and asking how they can design to induce public health. McCay discusses several factors that lead to increased stress for city dwellers, including socio-economic disparity, stimulation overload, a reduction in exercise routines, and connection to family and nature. He goes on the discuss the opportunity for urban designers to address these four key factors. This WG sees the same opportunities for the spatial web and we encourage AR creators to consider and design for these as well[150].

1) Open green spaces that are easily accessible and inclusive cut down on stress by providing a break from everyday demands. Spaces that are walkable and encourage social activity are most effective, but window gardens, and landscaped walkways can also provide benefit. Enhancing spaces in this way, that may not otherwise have these characteristics is an excellent use of AR.
2) Activity and exercise can help to ease symptoms of ADHD, dementia and even schizophrenia. Transport paths for biking and walking that are safe and convenient provide an alternate to sedentary lifestyles. Creative use of AR can expand the possibilities for activity in limited spaces.
3) Creating spaces that encourage social activity, with features that accommodate safe interaction such as benches, tables, grills, or even life size tic tac toe help to foster a sense of belonging and community. We can't well add furniture through AR, at least not soon, but AR is well suited for adding opportunities for social interaction.
4) Making a space safe and secure with good lighting, visibility, easily recognized landmarks and wayfinding. We already know that wayfinding is an excellent use of AR. Consider how adding these types of enhancements can increase safety and security of public spaces.

Similarly, architects and product designers use digital tools, but may not realize that could (and should!) apply their skills to designing an enhanced, digitally-driven-and-generated reality. Their knowledge about haptic and tactile feedback from objects and surfaces is invaluable.

[150] McCay, L., Designing good mental health into cities: the next frontier for urban design, in: Urban Design, Issue 142, Spring 2017, Urban Design Group, https://www.designcouncil.org.uk/news-opinion/designing-good-mental-health-cities-next-frontier-urban-design

Since the 1990s, research, experiments and activities at the intersection of digital media and architecture have been grouped under the term "media architecture". In 2009, the Media Architecture Institute[151] was founded as an international network. Establishing an exchange with product designers and the Media Architecture Institute would be a great advantage in the holistic design of the spatial web.

UI designers generally still focus on the design of 2D elements on framed screens. But stronger than architects, they already have to consider dynamic behavior and adaptability to individual users. And game designers even create complete environments for multiplayer use with complex interactions in lively environments. Even though there are usability challenges in game UIs, the benefits, affordances and efficiency of interactions are often not the central requirements for games, as it will be the case for most applications of the AR Cloud.

Experienced and open-minded experts from many areas are already members of this Working Group for user experience, accessibility and safety. To understand and master the possibilities and pitfalls of this new and all-encompassing design challenge of the spatial web, it will be imperative to actively engage experts from many design fields. Together we need to identify groups that can evolve and foster best-in-case user friendly AR Cloud experiences.

Current Status

Our Working Group's focus is on researching, reporting, and advancing design principles set forth in Augmented Reality technology with a focus on user experience, accessibility, and safety. We aim to learn from and partner with both industry and academia to accomplish this goal. As a not-for-profit entity the goal is to further advance the AR medium not for monetary purposes, but for the wider adoption of an AR Cloud framework that benefits the greater good. In the next section of this State of Augmented Reality report, we look at what the current state of designing for the best user experience in AR looks like and how it may evolve.

Lessons Learned from Existing Technology - The Importance of Research

Basic principles of user experience design still apply to AR. Although we must consider a variety of complexities, it remains fundamental to take a research based, user driven approach. Having clear, defined business objectives and functionality based on user needs and goals is good product design, but will also help to test the appropriateness of using augmented reality.

151 https://www.mediaarchitecture.org/

There is a tendency with a new medium to use it just because it is trendy. Therefore, it may be prudent to scrutinize each proposed use case to determine if there is a benefit to using AR that could not be achieved without it. Adding AR to an existing application or idea without a genuine motivation will almost guarantee a bad user experience. Keep in mind, at its core, AR should add a layer of value that improves the user experience and helps to reduce the complexity of simple tasks. Gratuitous use will do neither.

Augmented Reality is a shiny and new technology with high WOW factor. It's important to recognize how much this impacts user feedback. This can make it difficult to get good data and can even make it difficult for the designer and the researcher to uncover design flaws. Keen observation and listening is key. Discerning between the excitement over the technology and any obstacles to a natural user experience requires stern objectivity.,

Because the real world is an uncontrolled environment, research plays an even greater role in the design of real world augmentations. Through UX research methods such as task analysis and contextual inquiry, observing the user in their natural surroundings, we can learn a great deal can be learned about the user and the physical characteristics of the environment. We also have a responsibility to consider the effect of the augmentation on unintended users, or bystanders, which we can only really observe in situ.

The extensive augmentation of the landscape in the social AR game by Niantic, *Pokemon Go*, resulted in several types of disturbances to incidental bystanders. People not playing the game would suddenly find themselves overtaken by a mob of "hunters" in a particular area the game marked as a lure to capture characters. There were also incidents of trespassing on private property where people were again being drawn to uncover characters for capture. The viral nature of the game could not have been easily predicted from standard user testing. We require methods for uncovering these types of unintended consequences.

While conducting the contextual inquiry, it will be important to consider how various conditions of the environment might impact the augmented experience and how those conditions might change. Lighting conditions, connection speeds, furniture arrangements and other environmental factors such as temperature, humidity and noise will all impact the user's experience and the technology's ability to function as intended. Collecting as much data as possible will allow for a thorough evaluation.

By conducting a task analysis for each scenario, breaking it down and recording the steps required to accomplish the intended task will help to gain a thorough understanding of the user's perspective. Carefully thinking through interaction scenarios well before specifying technical requirements will help to ease unnecessary re-development efforts and make for a better user experience. Together, contextual inquiry and task analysis will provide the basis of empathy required for the design of a more natural experience[152].

While navigating the natural environment, there are many uncontrolled elements and unforeseen hindrances. Real world testing of actual task flows, with the target user, on the actual device in the intended environment is key to successful product development and design. It's also important to anticipate the users' needs and to consider the unexpected places the experience might occur in, besides the expected places.Testing in the wild is the only way to uncover the potential impacts and unintended consequences. Especially at this early stage of the technology, we have a responsibility to the public, to test as much as possible.

From Skeuomorphism to Direct Manipulation

Although there is much we can draw from the best practices of other mediums, augmented reality takes us out of the confinement of a screen and puts us back into the real world. Interacting with the real world through a device is not natural, so creating an interface that is seamless, unobtrusive and provides real value is needed. A great deal of research is taking place to develop more natural interfaces for interacting with the complexities of the real world. This is an area of tremendous opportunity.

[152] Miller, A., Things To Remember When Designing For Augmented Reality, UX Planet, June 20, 2017, https://uxplanet.org/things-to-remember-when-designing-for-augmented-reality-ecab17e3370e

The Mi.Mu gloves are a wearable musical instrument, for expressive creation, composition and performance and a wonderful example of the possibilities for a most natural interface. Source:@imgoenheap (Twitter)[153]

As with many technologies in the past, early AR cloud applications will simulate previously known interface expressions and mimic interaction paradigms from the analog world. Using familiar mental models allows people to get used to the new medium. Until the medium matures and the best means to exploit its benefits are fully understood, we rely on our understanding of existing mediums. Eventually, the adoption of these familiar forms are replaced just as the wire hinges on the early iPads have been replaced by one with flat information design. AR too will eventually find its own form that fits the medium and leverages its innate advantages.

The novelty of a new technology itself is often enough for users to get used to. Is it natural for them to want to user familiar interactions and it is helpful when designers ease users into a new technology this way. Look for ways to incorporate familiar gestures, similar to touch screens, with the hands in 3D such as squeezing, stretching and pushing objects directly. While it is common for people to use screen-based interactions such as pinch to zoom, swipe to see a new selection and tap to select, instead of simply walking closer to an object. While it's important to provide these touch-prompted interactions, it's also important to provide explicit

[153] https://www.youtube.com/watch?v=CvyVQqCO8pY

visual and audio cues to help people navigate the environment—teaching them how to interact in AR.

Also consider real world mental models for new types of interfaces such as lamp pulls, light switches, the drawing of curtains. These types of gestures engender certain expected reactions in the physical world. Digital objects need contextual reminders that proclaim they exist, have perceived function and that they can occupy both expected spaces as well as negative spaces across dimensions.

With mobile AR, the user splits their attention between the view of the real world seen through the device, the interface, and the real world itself. A delicate balance is necessary to keep the user engaged and also fully aware of their surroundings and it's important to avoid going back and forth between the device interface and the scene. Cheng[154] gives the example of a character that is talking. It is better to use an element within the screen, such as a speech bubble, rather than traditional subtitles at the bottom of the screen.

Considerable research is underway to develop more natural and accessible interfaces. In this section we looked at best practices for technologies that are readily available today. Using haptics, facial recognition, hand tracking, brain interfaces and galvanic skin responses are all covered in the *Direction of Evolution* subsection. Ethical consideration of some of these input modalities is covered in the *Ensuring User Safety* subsection.

Sensory Cues to Direct the User's Attention

People are accustomed to interacting with a flat screen. Because AR is still an unfamiliar medium, we need cues to encourage new users to fully explore the 3D environment. Consider visual cues such as movement to draw the eye in a particular direction. Appropriate use of flashing in the peripheral view can be especially helpful when an object of interest is outside of the preferred cone of vision.

The human cone of vision is very limited, but we can hear around us in 360 degrees. Audio cues are an excellent method to direct users' attention to off-screen objects. Spatial audio can also enhance the sense of presence and create a more immersive reality.

Other visual cues should help to describe the scene to the user. If an object is interactable, some cue should be given such as a glow, subtle animation, or sound. Maintaining a sense of depth is also important to merge the digital with the real

[154] Cheng, J., Early Challenges in AR UX, Google, March 8, 2018, https://design.google/library/augmented-reality-ux-design/

convincingly. Because text or symbols have no inherent sense of scale, it helps to use a character guide to convey depth accurately in instances such as wayfinding.

Lighting can enhance the feeling of presence. In an ideal situation, sensors would detect the direction and quality of light hitting a scene to replicate it for the virtual object, but more practically an approximation can be used. The simplest solution is to use high-noon sun directly above the object. This will interfere the least with most actual lighting situations. Adding shadows will help to give an object weight and add a sense of depth. This is especially important for mobile AR because the perception of depth relies on binocular vision and a phone screen flattens the appearance of objects in 3D space[155].

When designing Animoji type experiences (e.g. Snapchat filters), provide guidelines for user tracking. Using facial expressions to add unexpected additional augmented features such as a mustache when detecting a mouth opening, changing eye color on a wink, or alighting a hummingbird on a perched finger can bring a sense of added delight to the user.

Spatial Typography

Dong Yoon Park is an expert of 3D typography and has created several experiences for the HoloLens that explore text in mixed reality. In an article for Medium, *Type In Space—Explore Spatial Typography In Mixed Reality with HoloLens*[156], he puts forth a list of recommendations for working with text in virtual and augmented spaces, which is included below.

Basic visual design principles still apply and are especially important to follow for rendering text. Use high contrast lettering, dark over light, light on dark or complementary colors. Text cards can help immensely to simplify the noise in the background such as varying light conditions and textures. Be prudent when selecting fonts and keep them simple. Text is inherently a 2D medium and several challenges present themselves when rendering in 3D.

Chris McKenzie and Adam Glazier from Google presented a system for text scaling at Google I/O in 2017. They based the system on an angular unit called a dmm (pronounced “dim”), or *distance-independent millimeter*, that maintains the relative size of the text at any distance being viewed. Using the dmm, they developed

[155] Wilson, T., The principles of good UX for Augmented Reality, UX Collective, November 13, 2017, https://uxdesign.cc/the-principles-of-good-user-experience-design-for-augmented-reality-d8e22777aabd
[156] Yoon Park, D., Type In Space—Explore Spatial Typography In Mixed Reality with HoloLens, July 18, 2018, https://medium.com/@dongyoonpark/type-in-space-explore-spatial-typography-in-mixed-reality-with-hololens-28e1942ba2cf

interface guidelines for text and button sizing. They have also developed a system for implementing the concept in Unity. The resource can be found in Google's developer documentation under "Resources"[157].

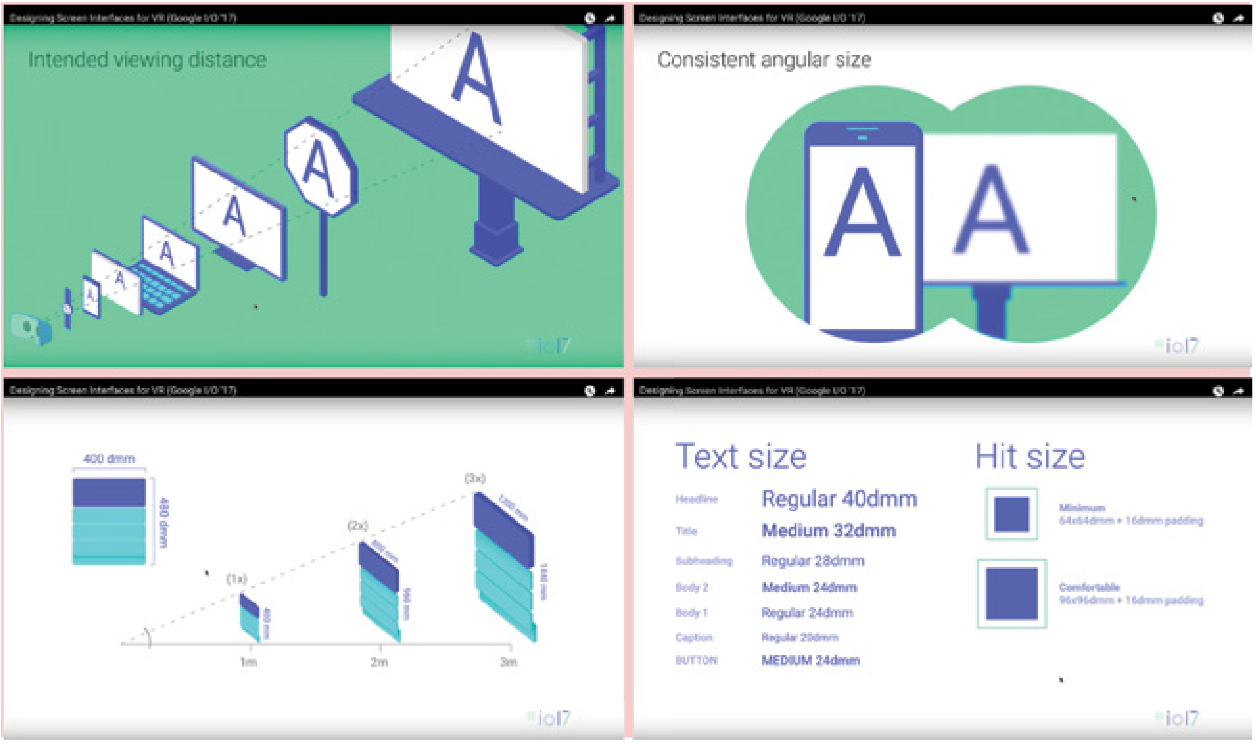


Image excerpts composited from Google I/O 2017. Source: McKenzie and Glazier

Although quickly improving, display resolution is still an issue for many consumer grade hardware manufacturers. Often, we can see an artifact known as the "screen door effect" on text rendered in 3D and we should take care to avoid it. Unity has developed a great solution using SDF technology that can render crisp outlines of the text regardless of the distance from the viewer.

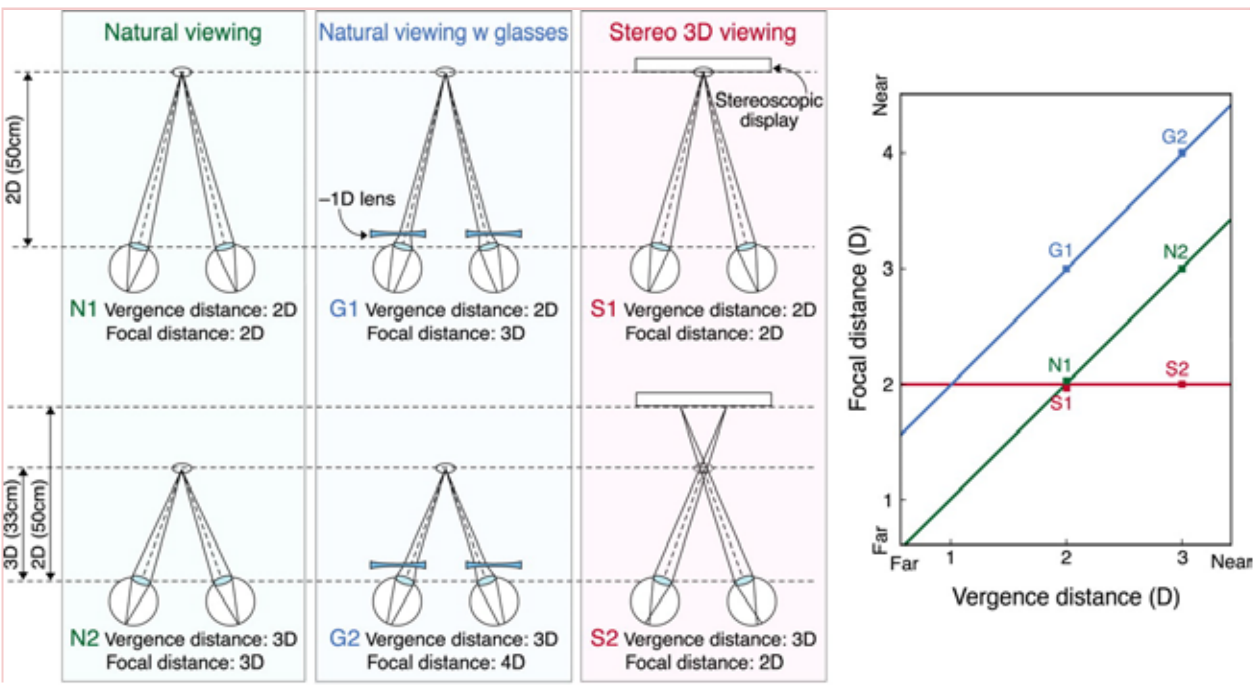
Extruding text to make it 3D hurts with the recognition of the letter's silhouette, and can degrade legibility, so should remain flat. For especially wide screens, a slight curvature of the screen helps to reduce skewing of the text at the edges[158]. Likewise, text should remain facing the viewer, if not, the skew will distort readability. A technique known as "billboarding" allows an object to rotate freely in space, so that it may always face the user by orienting to the user's line of sight.

---

[157] https://developers.google.com/live

[158] Alger, M., Visual Design Methods for Virtual Reality, MA Moving Image, September, 2015 http://aperturesciencellc.com/vr/VisualDesignMethodsforVR_MikeAlger.pdf

Due to the vergence-accommodation conflict, optimal viewing distances for virtual objects is between 1.25 meters and 5 meters, with many head-mounted displays optimized for a 2 meter focal point. In most headsets, users will always accommodate, shift focus by adjusting the aperture of the pupil, to the focal distance of the display (to get a sharp image), but converge, direct the rotation of the eyes to meet at the object to maintain binocular vision, to the distance of the object of interest (to get a single image). Therefore, we recommend locating large amounts of text and user interfaces at the focal point for the headset, about 2 meters. This allows the user to accommodate and converge at the same distances naturally. It is best to locate all objects of focus within this range. When the natural link between convergence and accommodation is broken, it can lead to visual discomfort or fatigue.



For a fixed focal distance, such as in a head mounted display or optical glasses, the natural vergence distance and focal distance intersect at the focal distance of the device's display. Source: Shibata[159]

Ensuring User's Safety - The Physical Environment

Besides privacy and security issues, ensuring the user's physical safety is of utmost concern with AR. There are a few considerations when trying to determine just how much mediation to require. What is enough to keep the user engaged and find value, and what is too much that may cause distractions and be potentially unsafe?

[159] Shibata; Kim; Hoffman; Banks, The zone of comfort: Predicting visual discomfort with stereo displays

Luckily, fundamental design and user experience principles still apply. Keeping the augmentation simple, adding only what we need to achieve the goal will help to avoid distracting experiences. It's important to use restraint to avoid interference and minimize cognitive load. Animation can help direct the attention when needed, but we should take care to abstain from overstimulation. When NFL broadcasters initially added the first down line drawn overtop of the live video feed, it was useful. After they started drawing large animating arrows following the plays it became distracting and took away from the experience.

Maintaining a narrow depth of field helps the user in several ways. Placing objects beyond the user's field of view flattens the scene and makes it difficult for the user to perceive depth, which may lead to confusion and potential accidents. By limiting the augmentations to only what is in the user's field of view, the user can more naturally make sense of the relationship of the virtual objects to the surrounding environment. Prioritization of reality helps ensure user safety. We consider a generally acceptable depth of field as between 1 and 5 meters. This is due to the vergence accommodation discussed in the *Spatial Typography* section above.

Keeping the line of sight clear of obstruction will help to keep the user oriented in space and aware of his or her surroundings. An excellent example of maintaining the users' lines of sight is the Land Rover Bonnet example mentioned in the *Companies Doing Work in this Field* section below. The Pokemon Go experience, which we discuss in depth in the next section, required users to hold the device up and use their finger to catch the AR critters. This obscured the line of sight, distracted players from their environment and caused a number issues for themselves and those around them.

The Unintended Consequences of Pokemon Go

With the best intentions, the makers of Pokemon Go designed the game to get people off of their screens and out into the real world. Little did they realize what a craze would ensue. They even had the forethought to use the vibration of the phone to notify users when they were near a character to avoid people being distracted by the game and lose attention to their physical surroundings. Those efforts made little difference in the face of people's unabashed obsession over the game. The novelty of the technology, the social and real world interaction compelled people of all ages to abandon reason and sometimes risk personal safety.

Besides traffic incidents involving both distracted drivers and meandering pedestrians, people cut and bruised themselves by tripping, falling and banging into things while playing Pokemon Go. Players reported other, more lurid dangers.

To help your friends catch more creatures, a player may set up a public "lure" that attracts characters to a public "Pokestop" exposing the characters to capture. Players lure others to the location—sometimes in droves, which has also created issues for property owners. Of greater concern is the use of the lure feature for unethical purposes. Incidents of robbery, carjacking, trespassing and even some cases of sexual assault took place with perpetrators getting convicted.

It is often difficult to predict unintended consequences with the use of a new technology. We never know for sure how a feature or application will ultimately be of use until released into the wild—even with user testing. With augmented reality and emerging technologies, the novelty alone is enough to draw people's attention away from their physical surroundings, even when the attempted avoidance of distraction was a consideration in the design.

With new technologies, comes increased responsibility for the creators. Although the creators of Pokemon Go understood the concern for distraction, they didn't play the scenarios out far enough to design for the safety issues that ensued. Perhaps an added step to the design process needs to be hypothesizing about unintended consequences, dreaming up worst-case scenarios and doing our best to design safety precautions into the experience.

Accessibility and Inclusivity

Rather than an afterthought, or worse yet, omission, we have shown designing for accessibility — adding improvements to be more inclusive of persons with disabilities — makes for a better product overall, for all. Designing for accessibility in AR is not that much different from designing for accessibility in other mediums. What really stands out is how much we can use AR technology to improve the lives of the disabled and create conditions that allow them to more fully participate, prosper, and reach their full potential.

Creating accessible and inclusive experiences isn't just a nice thing to do, from a societal view, it's crucial to economic sustainability. Leading world organizations such as The Organisation for Economic Co-operation and Development (OECD) and the International Monetary Fund (IMF) have concluded that deepening inequality leads to a decline in economic growth. When support is provided in the areas of greatest need, circumstances are created that allow those who have been left behind to participate and contribute fully. By increasing accessibility, we move towards everyone being treated equitably, and build a future based on inclusion and a society based on true representation[160].

[160] Blackwell A. G., The Curb-Cut Effect, Stanford Social Innovation Review, Winter 2017 https://ssir.org/articles/entry/the_curb_cut_effect

There is a statute in the European Union called the Movement of Freedom, where all citizens have the right to move freely between states; to work, play, study and live freely between nations. Yet persons with disabilities cannot enjoy their freedom of movement. As in all nations, the disabled have difficulties accessing mainstream goods and services. They do not have the same opportunities as persons without disabilities: they face barriers when studying, looking for a job, travelling, purchasing goods and services or accessing information.

Barriers to access go beyond what we typically think of regarding physical or environmental limitations. We need to consider institutional barriers also: laws, policies or practices that are restrictive or discriminate against people with disabilities and even attitudinal barriers that may be dismissive, discriminatory, fearful or cruel. As builders of the future, it is imperative we are mindful not to create new barriers and to look for ways the technology can remove existing ones. When we remove barriers, people with disabilities can be independent and equal in society. There is an opportunity to build inclusivity into the very fabric of the AR Cloud.

Although AR poses new opportunities for inclusive design, there are still basic protocols we can adopt from web and mobile development. When doing initial user research and subsequent usability testing, including users with a variety of disability types is key. As we will see, the definition of disability needs to extend beyond what we classically consider, such as persons in a wheelchair or with visual impairment, and include temporary disabilities such as a broken arm or fatigue, age related and even psychological impairments.

To get a better understanding, it is important to observe people while using assistive technologies, such as screen readers, and voice over controls. In fact, to evoke empathy within yourself, try navigating with an array of assistive technologies —try navigating a web page using only your keyboard, learn how to use a screen reader such as VoiceOver on your Mac, or adjust your accessibility settings to experience a site as if color blind. Consider the field of view; many sight challenged persons use some form of a magnifying glass. Try looking through a narrow opening made with your fist to mimic the restricted field of view of a magnifier and determine what content hides on an initial vertical scan. These experiments will give you firsthand knowledge about what it’s like for many of your users so you can design to include them[161].

[161] United States Government, Accessibility for Teams - A ‘quick-start’ guide for embedding accessibility and inclusive design practices into your team’s workflow,https://accessibility.digital.gov

"...Almost everyone experiences some kind of disability either permanently, temporarily, or situationally. For example: having only one arm is a permanent condition, having an arm in a cast is temporary, and holding a baby in one arm is situational – but in each case you're restricted to completing tasks with one arm."

Whitney Queensbery wrote an article[162] that is a wonderful reference covering a wide array of impairments.

From Web 2.0 to Spatial Web

The same principles that apply to designing accessibility for the web will apply to AR, but we may need to create new semantic markers. Consider that the "screen reader" will need to become the "scene reader". We will need similar descriptive elements such as the landmark elements of HTML and semantic markers of ARIA to describe the entire world, not just a page. We will need to ability manage focus and skip the navigation to the core experience resembling the flow of a screen reader. Similar to well-defined image tags, we will need object tags that define not only visual descriptors but maybe size, location and temporality. New specifications will be necessary for touch targets and accommodations for reaching a target will be a requirement in the 3D spatial web. Owlchemy Labs[163] is a game company that is using suggestions such as these to improve the usability of their games. To make it easier to grasp objects slightly out of reach, the object responds to gestures that are "close by" and will move within the player's reach. It is a great example of using accessibility directives to make a better experience for all. Other specifications will be necessary for eye tracking, hand tracking, tapping, voice control, and new forms of haptic and neurological input such as reading brainwaves, EEG and EKG.

Of great promise is the gaze, gesture, and multimodal input research of Mark Billinghurst[164] . In the true story *The Diving Bell and the Butterfly***,** journalist Jean-Dominique Bauby suffered a massive stroke that left him with locked-in syndrome—a condition in which a patient's mind is fully functioning but cannot move or communicate verbally due to complete paralysis of nearly all voluntary muscles in the body except for eye and slight head movement. Gaze as a singular input is difficult to use—the eyes are continually scanning the environment so the user can not select everything they look at, but Billinghurst's continued work with natural eye movement interactions could unlock these patients and free them to communicate with the outer world.

[162] Queensbury, Whitney, Personas for Accessible UX, Rsenfeld Media
[163] https://owlchemylabs.com/
[164] https://medium.com/@marknb00/multimodal-mixed-reality-b4ab8ddc304b

By combining input modalities, the limitation of one can offset the limitations of the other. Billinghurst found that combining gesture or head pointing with eye gaze, was nearly ten times more accurate that gaze alone. When combining gesture for qualitative input such as furniture arrangement coupled with speech for quantitative input such as precise measurements, accuracy was higher and arrangements reached completion 30% faster. Users report these types of interfaces provide a higher level of usability and natural interaction.

Borrowing also from designing for web accessibility, how will designing for the spatial web let users know the nature of an experience a user can link to, prior to accessing the link? What could the nature of a link be in 3 dimensions? Captions, transcripts, and other supportive data for images, videos and data can take advantage of spatiality. Consider a data visualization that has not only text captions but also directive symbols guiding a user to experience the visualization spatially and temporally for the deaf. Or consider auditory transcripts that take advantage of spatial audio to give the blind or sight impaired a sense of scale, location, direction and movement of the data. With AR and the Spatial Web it is truly a multisensory experience where all 5 senses are taken advantage of in new and promising ways, not just for the disabled, but for all users.

Direction of Evolution

Augmented Reality is becoming more adopted every day. And with this technology becoming more accessible and present in the world, it will come down to well-designed experiences for wide acceptance. Our Working Group's goals as outlined above is to highlight and be a champion for the power of well-designed experiences. We as developers and consumers also need to know of society's fears towards new technology and how we can use this game-changer for exploitative means or financial gains. We believe in a free and open AR platform. For us to get to the wide acceptance of such a platform, we as the developers and consumers plan on studying both good and bad design and uses of augmented reality experiences. In this subsection, we outline the human psyche's usage of mixed reality experiences, how they are designed, and current issues facing the technology with wide adoption.

One, Yet Individual Reality

A possible overarching "mantra" for integrating disciplines stems from constructivism and perceptual psychology: users perceive digital and analogue representations and functions as a single integrated reality—the boundaries between digital and analog, virtual and physical blur. For the individual, only the meaning of the information or services in the current context is relevant - no matter how it is provided. What the individual internalizes as "their reality" is subjective and potentially different from others do synthesize from the same environment.

As applications dynamically adapt to the individual way of perceiving and understanding, they are more efficient as a medium and tool for the individual. However, personalization assumes that the technical platform has profile information for each user, which requires the proper handling of privacy issues.

The notion of different realities raises difficult philosophical questions and underlying psychological and social phenomena that need more exploration. In a multi-dimensional reality composed of multiple analog and digital layers from miscellaneous sources, which reality layer is the one to believe?

Looking at today's social media news feeds that deliver us highly charged messages according to our preferences has not only positive but also negative implications, as the messages about what is real, even in our understanding, become highly subjective.

Hyperreality and its Relationship to Augmented Reality and OARC

Coined in 1981 by French sociologist Jean Baudrillard[165] , "Hyperreality" is a postmodern belief that as society advances technologically and that technology molds around the human world, individuals will not have a clear distinction between what is the real world and what is a fictionalized version of the world thru the lens technology offers. It's a dystopian fear that humanity references and discusses we develop more technological advancements like virtual reality or augmented reality. It is the fear that humanity may embrace VR and AR in ways beyond what people consider the "natural order" of how humans interact with the world, such as not needing to leave your home ever again as referenced in the novels "Ready-Player-One" and "Snowcrash". Hyperreality warns of humanity's escapism desires[166] - a choice to accept fabricated realities as believed to be true reality. It's also a very dystopian outlook believing that humanity will live in a technologically created reality that hides the true nature of the world.

The fears of Hyperreality may never go away. Anytime a radical new technology tool becomes integrated into everyday life, humans can be hesitant to embrace fully out of fear of being forever changed by machines. Filmmaker and designer, Keiichi Matsuda premiered a 360-video film, *Merger*, that dealt specifically with these fears

[165] Baudrillard, J., The Precession of Simulacra, Simulacra and Simulation, 1981, Translated by Sheila Glaser, University of Michigan Press, 1994
[166] Hall, A. C. O., "I am Trying to Believe": Dystopia as Utopia in the Year Zero Alternate Reality Game, Eludamos. Journal for Computer Game Culture, 2009, p. 69-82

of Hyperreality and AR's power on productivity. "We need to restructure our society in a more radical way, where automation becomes an asset instead of a threat," says Matsuda[167].

The OpenARCloud community's Working Group on User Experience, Accessibility, and Safety, believes that Augmented Reality should be a tool to enhance our world and our storytelling abilities, but not be a replacement of the physical world as we know it. Augmented Reality is a growing new frontier with vast possibilities for Hyperreality dangers, but OARC believes more in the positive benefits of this technology and how we can use it as a tool to enhance aspects of our lives, not replace them.

The Role of Biometric Data in the Spatial Web

We all know that data is being collected everywhere. With the impending AR Cloud and proliferation of wearables, it's hard to imagine, but data capture will increase exponentially. And you can bet, it's personal! Cameras will be everywhere, every person with a device, wearing a watch, a pair of glasses, even jewelry will, will be a source of capture and analysis. These sensors will determine not only where you are and the characteristics of your surroundings, but details about your health, your thoughts, your feelings and even how strongly you are feeling them. How this interacts with system architecture and user experience is in tune with Working Group's core tenant on User Safety. AR Technology should be a benefit to society and not something to fear in the hyperreality states mentioned earlier.

Headsets and glasses can track your eyes to determine what path your gaze takes and what has visual salience to you. By attaching electroencephalography (EEG) sensors, we can measure cognitive load, and show an aversion to a task, its level of challenge, and levels of attentiveness. We can correlate anger, surprise, fear, joy, sadness, contempt and disgust with certain muscles in the face and can read them by detecting electrical impulses in the facial muscles through electromyographic (EMG) signals. Although the hardware required to do this is too bulky to be practically placed on the face at scale, computer vision is becoming sophisticated enough to read emotion through body positioning, facial expressions can't be that far off.

Often increased security comes with a decrease in ease of use, but it doesn't have to. This Working Group's focus on quality design can aid in the study and dissemination of findings based on user experiences. Security systems that are inconvenient may

[167] Winston, A., Keiichi Matsuda explores dystopian future of the workplace in new film "Merger", de zeen https://www.dezeen.com/2019/01/17/merger-keiichi-matsuda-future-workplace-augmented-reality-video/

not be more secure if people resist using them. People may not comply with systems they see as intrusive, restrictive or a nuisance, leaving themselves and the system vulnerable. Security tools and software must be hassle free and integrate smoothly into a user's workflow. System prompts and messaging need to be not only technically correct but also from the user's perspective, friendly and information—helping the user to know what action to take. Choices should be clear, within the context and the consequences known before taking the action. People want their devices and system to be secure and will comply when a hassle-free experience is available[168].

As discussed earlier, many companies are working to evolve this technology ethically for the good of humankind. There are many positive uses of this data and hence the real reason for collecting it. We can use it to monitor our health, both physical and mental. This allows us to optimize our health we can be happier, more productive and increase our overall satisfaction with life. All this data can provide VIP services and experiences that are move convenient, and with less friction. Services such as having a rental car meet us at the airport without lines to wait in or papers to sign. It will allow us to do away with passwords forever, but identity theft, although far more rare, becomes much more gruesome.

Design on the Brain: Neuroarchitecture

An online-article by Winer and Keim[169] introduces "neuroarchitecture" as a field that "applies the scientific rigor of neuroscience to the world of design, seeking to better understand how humans perceive, experience and react to (stimuli from) built spaces at a biological level." The historical story of Jonas Salk who discovered the cure for polio not in his lab but at a Franciscan monastery as a "quiet abode where he cleared his mind," illustrates the underlying relationship between the mind and architecture. The aim "is to extract findings that can be practically applied, lending a scientific backing to designing environments that positively shape the user experience, namely human health and happiness."

Technologies like human monitoring techniques combined with VR already allows researchers to "observe brain activity as people interact with components of the built environment". "When combined with physiological data such as heart rate,

---

[168] Strehlow, R.,Cyber Security Requires an Important Ingredient: Strong UX, 2018, https://hackernoon.com/cyber-security-requires-an-important-ingredient-strong-ux-d0727a0c076

[169] Winer E., Keim J., Design on the brain: Combining neuroscience and architecture, International WELL Building Institute, October 9, 2018 https://www.wellcertified.com/en/articles/design-brain-combining-neuroscience-and-architecture

these findings reveal information about a person's mental state, stress levels and learning mechanisms while interacting with the space."

"Neuroarchitecture" demands interdisciplinary collaboration to succeed and should apply to real-world projects to translate "science into evidence-based practice". "Data used to inform future design will no longer need to be obtained retrospectively from [...] post-occupancy surveys". It would be possible to "proactively combine information about a building's context and intended function with knowledge about the human mind to maximize user benefits and well-being". Dr. Eve Edelstein, a neuroscientist, architect, anthropologist transcribes this as "laboratories of human experience."

As science leans more on the power of Augmented Reality as a tool at their research disposal, this Working Group plans on being a champion for its benefits. Through prototypes of our own, partnerships with research teams using AR technology, and publishing of this work - our goal is to make the adoption and ease of use with Augmented Reality for studying human interactions more prominent.

Companies Doing Work in this Field

As Augmented Reality technology is in its infancy, countless offshoots of AR concepts are being developed for the first time or expanded upon. OARC plans to learn from and share best practices of as many of these experiences as possible. We believe that AR technology focused on User Experience, Accessibility, and Safety can be divided into the following categories:

- Visual Augmentation - the layering of information onto a user's sight for an AR driven experience
- Aural Augmentation - the layering of information through sound for an AR driven experience
- Haptic Augmentation - the layering of information through touch and vibrations for an AR driven experience
- Assistive AR Technology - Technology that uses a variety of human senses to enhance, aid, or serve in place of another sense in a technology-driven experience
- Self-Reliant AR Technology - Technology that uses computer processing power to understand the world around it through image processing for automation of a task

Visual Augmentation

The most common usage of Augmented Reality is Visual Augmentation. These AR experiences work through smart phones, cameras, or wearable headsets by layering information on top of the real world through the lens. This technology includes being able to do everything from the practical to the experimental. Today, you can use Visual AR technology to test new eyewear, apply a digital tattoo to your body, repair a rocket engine and even understand foreign languages in real time.

One of the earliest and widest used examples of Visual Augmentation is the social media platform Snapchat. Snapchat users share photos and videos between their social network and have the ability through the app's image recognition and AR ability to add additional elements on top of what gets streamed out to their network. With the touch of a button, the application identifies a user's face and can attach 3D animal ears or add a 3D layer of makeup that moves with the subject's head. The wildly popular social network introduced millions to the power of Visual Augmentation by making it easy to alter their world through digital means.

A powerful aspect of AR as a platform comes from asking "What if?" and allowing a user to explore those questions by altering the digital representation of the real world. Another example of Visual Augmentation exists through the tattoo focused app, Inkhunter[170]. People considering getting a tattoo on their body can choose designs and scan the location on their body they hope to get the ink. The app will show the user a representation of what the tattoo can look like on their skin in that location before deciding on a permanent choice. By leveraging the power of "What if?", Inkhunter demonstrates to the tattoo community how AR can benefit its users.

Google Lens is the Google's Visual Augmentation tool that allows a user to interact with the world in a variety of ways from identifying buildings through image recognition and GPS data to pulling up information on where to purchase a specific object scanned with a user's camera lens. One of the more powerful examples of Visual Augmentation through Google Lens is the ability to translate languages in real time using a smartphone's camera[171]. A user holds the phone up to written text (a sign, paper, ticket, etc.) and Google Lens alters the image to display that same text in the user's language of choosing. This technology is connecting the world and breaking language barriers.

Visual Augmentation capabilities also can help us enhance the world as we move contextually through our day. We can look to examples that are not necessarily user-initiated like in the case of autonomous vehicles that are replicating the visual experience by technologies like LIDAR, Radar, and 360 cameras. Companies using and investing this technology include but are not limited to, BMW, Audi, Tesla, and

[170] Morgan, B. 10 Fresh Examples Of Customer Experience Innovation, Forbes, April 15, 2019. https://www.forbes.com/sites/blakemorgan/2019/04/15/10-fresh-examples-of-customer-experience-innovation

[171] https://techcrunch.com/2019/05/07/google-lens-can-translate-foreign-languages-in-photos-and-read-the-text-back-to-you/?guccounter=1

Jaguar Land Rover with their Transparent Bonnet concepts.. This technology allows for visual representation and augmentation to display new interpretations of complex sensor based data.

The Transparent Bonnet *uses a sensing camera to detect the road conditions and a second camera to project that image onto the bonnet, or hood, to assist the driver in difficult conditions ahead. Here augmented reality is being used to essentially render the hood and engine invisible. Source: Jaguar Land Rover*[172]

Aural Augmentation

When people think of Augmented Reality experiences, their first thought is looking through a lens that enhances the world view with new information. Aural Augmentation is an offshoot example of AR's potential for a more powerful future that does not use visual cues at all, and is entirely audio driven for the user based on data provided to the experience. The Marvel Cinematic Universe has shows Iron Man making calls to his personal assistant for "What am I looking at?" to "Best route to where I need to fly?" and that's very close to the idea of Aural Augmentation. The idea is not entirely new, as GPS systems have been providing driver's with audio cues on when to make appropriate turns to a destination for years. The idea of

[172] https://www.landrover.com/experiences/news/the-transparent-bonnet.html

Augmented Reality without sight can be a powerful tool. Using GPS data alone, AR experiences can funnel a user lots of data to guide them, instruct them, or remind them - becoming a useful tool. There are companies looking to build upon humanity's relationship with sound for gaining more information through technology rather than visual enhancers.

The audio company, Bose[173], has joined the AR development world bringing it's knowledge of crafting high-end speakers and marrying that technology with lightware and stylish eyewear frames and AR frameworks. Bose's goal with Aural AR is to give users audio cues based on where their heads may look for unique new experiences. Examples have included golfer's getting beeps for instruction on course grading based on where they are looking on a course (tied to their GPS location) and guided meditation sessions (based on head movements while wearing Bose AR glasses.)

Amazon is also experimenting with Aural Augmentation on the go. The Alexa personal assistant has led to mass adoption of virtual assistants in home usage from getting weather reports, ordering laundry detergent, or controlling smart home devices by voice cues. AI Voice Assistants can be considered Aural AR experiences when relying solely on audio cues to relay information. With the upcoming Alexa-powered earbuds[174], Amazon hopes to allow users to make information calls based on their GPS location and data from other areas in the world and deliver that information back to the user with the familiar assistant voice.

As another feature to Google's powerful AI assistant, Google's Pixel Bud headphones[175] (and any headset with Google Assistant capability) bring visions of a science fiction future into present day by offering a way to translate foreign languages in real time. Through Aural Augmentation, barriers like language discrepancies can fade away, bringing humanity closer together. This technology is showcasing that Augmented Reality can be used interchangable for any tech experience that is enhancing the real world for the user.

Haptic Augmentation

While Visual and Aural Augmentation enhance the world through sight and sound, Augmented Reality experiences still are separate from the real world in that users cannot touch or feel holograms and virtually placed elements. Haptic Augmentation

173 https://www.mobilemarketer.com/news/bose-brings-audio-ar-shades-and-festival-content-to-coachella/552302/
174 https://www.retailwire.com/discussion/will-alexa-earbuds-advance-amazons-virtual-assistant-ambitions/
175 https://www.engadget.com/2017/10/04/google-pixel-buds-translation-change-the-world/

and experiences is a budding field focused on bridging the gap of what is the real world and what is the enhanced world.

The idea behind Haptic Augmentation is to add the sense of touch to an experience. Despite haptics are still in early stages of development, there are several companies making significant advances in this exciting area. Our sense of touch is how we interact with objects in the real world and will allow for a more natural interface with the digital and virtual worlds. As these products hit the market, they will establish an entirely new way for humans and machines to communicate and change the way we interact forever[176].

Several experimental tangible interfaces showed that a thoughtful combination of digital technologies with physical setups can be synergistic. Ishii et al.[177] found that city planners traditionally need to integrate at least three different forms of representation (sketches, physical models, computer simulation) of spatial concepts and ideas into a single mental construct, distracting them from the central design process. Their proposed AR-Workbench "Luminous Table" integrates physical and digital representations and objects into a single layered interactive information space. Using "tangibles" and "phicons", such as wireframe and plexi-models or a handheld camera symbol, allows natural user collaboration and intuitive sunlight and traffic analysis with direct, non-symbolic interaction. Even this table-like setup does not support depth perception and has limited immersion, these techniques democratize knowledge by facilitating the processing of complex information for citizens.

Teslasuit[178], a full body haptic suit lets a user touch and feel a virtual reality environment. This revolutionary advancement in immersive experiences won a coveted innovation honoree award at CES 2019 for its unique approach to making the human mind “feel” like what they are seeing is actually happening. Features of the suit included full body motion capture, climate control, and biometry systems.

On a smaller scale than full-body coverage smart clothing, HaptX[179] offers wearable gloves to pair with Virtual and Augmented Reality experiences. The gloves allow a user to feel the shape, texture, weight, and motion of virtual objects with no perceivable latency in the feedback. With 130 unique points of feedback that

[176] https://www.ultrahaptics.com/news/blog/ces-2019-review/
[177] Ishii et al. Augmented Urban Planning Workbench: Overlaying Drawings, Physical Models and Digital Simulation. In Proceedings of the 1st International Symposium on Mixed and Augmented Reality (ISMAR ’02, 203-). Washington, DC, USA: IEEE Computer Society. https://dl.acm.org/citation.cfm?id=854980
[178] https://teslasuit.io/
[179] https://haptx.com/

displace the skin up to 2mm, HaptX opens up possibilities in experiencing with virtual objects in exciting new ways.

Deduced from these experiments, AR Cloud applications can unite (existing or specifically created) physical objects with digital spatial content and user interface elements not only to improve ease of access and comprehensibility, but to amplify a sense of “real”.

Assistive AR Technology

One of the main focal points of this Working Group’s mission statement is in the study, advancement, and development of assistive technology and that technology’s relation to an AR enhanced world. We can look at assistive AR Technology as any technology with a focus on aiding a user who may not have all the physical ability a majority of users may have, and allowing that user to engage with an experience in the best way possible. WG8 believes the power of the AR Cloud should be for everyone and we plan on being champions of assistive technology to make that a reality.

There are several technologies commercially available focused on assistive technology.

Tecla[180], created by Komodo OpenLab, gives individuals with physical disabilities the ability to communicate, control, and connect with the world through programmable smart switches and triggers. A user without complete control of their arms can use this technology to access other technology and the internet just as able-bodied users can. Tecla is for anyone who can’t easily use a smartphone, tablet or computer. This includes those with limited upper-body mobility resulting from spinal cord injuries, multiple sclerosis, ALS, muscular dystrophy, cerebral palsy, brain injuries, or stroke.

The vOICe vision technology for the blind[181] is another great example of technology augmenting the world for its users. The vOICe takes live camera views through image-to-sound renderings and feeds this back to a user in real time. Images convert into sound by scanning them from left to right while associating elevation with pitch and brightness with loudness. With typically an hour of training, vision impaired users have been able to identify objects, walk unassisted down hallways, and more with the skill level of someone with 20/250 vision[182].

---

[180] https://gettecla.com/

[181] https://www.seeingwithsound.com/

[182] https://www.theguardian.com/society/2014/dec/07/voice-soundscape-headsets-allow-blind-see

There have been some recent efforts to approach AR experiences from an accessibility-first perspective. These efforts take a human disability as a focus of design and end up making an application useful for most people. We can see this approach in the Soundscape project from Microsoft[183]. Soundscape was built for people with visual impairments as a tool to help them understand their surroundings by describing what is around them using spatial audio cues. It's backed by a powerful map system that contains rich data about points-of-interest nearby. When we take approaches like the one described above when building applications for augmented reality, the work opens up new design paradigms that are applicable to most people. Especially, when those people have the same set of senses available to them.

Self-Reliant AR Technology

Augmented Reality's power to our daily lives can not only be seen in examples as a traditional tool, but as an entire new subset of automation. Through image processing and augmented reality frameworks, we teach computers to "see" as humans do and "understand" the world as it exists. This next level of self-reliant technology that can take over a series of tasks for humans through automation and image processing can lead to wide adoption of AR through its ease of use alone. We find the most present example of self-reliant AR technology in vehicle automation and self-driving cars.

While most of this Working Group's focus will be on the development and study of Augmented Reality on a small scale with handheld devices or wearables, the technology is about to change the landscape of everyday life with the adoption of AR assisted driving and the future of automotive.

A wide variety of use cases already have been studied, including increasing older drivers' concentration and situation awareness with head-up AR displays[184], or supporting during driving phone calls with collaborative AR[185]. Other studies focus on using AR in automated vehicles to improve the passengers' acceptance and trust of the systems decisions by making them aware of its choices' reasons[186]. The

[183] https://www.microsoft.com/en-us/research/product/soundscape/

[184] Pampel et al., An Investigation of the Effects of Driver Age When Using Novel Navigation Systems in a Head-Up Display, in Presence, vol. 27, no. 1, pp. 32-45, March 2019.

[185] Kun, et al., Calling while Driving Using Augmented Reality: Blessing or Curse?, in Presence, vol. 27, no. 1, pp. 1-14, March 2019

[186] Wintersberger, et al. Fostering User Acceptance and Trust in Fully Automated Vehicles: Evaluating the Potential of Augmented Reality, in Presence, vol. 27, no. 1, pp. 46-62, March 2019.

diversity of these use cases point out that future automobiles will serve as AR devices, just as smartphones or smart glasses.

In this context, vehicles will be potential users of the AR Cloud. Thus, further work will be necessary to measure how certification authorities and end-users trust data coming from the AR Cloud to be used in such critical systems. However, some lessons learned in experimenting AR for automotive already can be generalized to mobile AR. For example, the results of the study presented by Pampel[187] allow to conclude that the display position have an influence on the user's awareness of its environment, and by the way on its security. The users of the study where better aware of the situation when using head-up displays rather than head-down displays. This shows that designers of mobile AR experiences should prefer displaying artifacts at a position that keeps the user's head up (e.g. in the far field, or at the head's height).

As explained by Wintersberger[188], one of the critical problems in automated vehicles is the lack of user trust in the vehicle's decision. The hypothesis is that AR could be a way to increase the understanding of the taken decisions. To show this, the authors proceeded two experimentations in a simulator. In the first one, they tested an artificial full-sized windshield display to augment traffic objects (e.g. other vehicles) outside the vehicle. The augmentations generated were colored triangle annotations overlaying the upcoming traffic objects. The color of the annotations depends on overtaking distance for oncoming vehicles, and on security distances for preceding vehicles. In this study, the augmentations were generated using data coming from sensor fusion and C2X communication which used by the car to avoid obstacles and control the speed.

The recent progress of these systems now allow automated vehicles to drive faster in a dense fog environment than in manual driving. If it represents an improvement in autonomous vehicles decision speed, it can be more stressful for the user and reduce its acceptance of technology without augmentations. Test scenarios, with and without traffic objects augmentations, have proceeded involving 26 student-volunteers. Afterward, a 20 minutes interviews of each participant took place to let them compare the scenarios with 7-point Likert scales. As a result, the authors found that traffic objects augmentations led to a significant improvement in user trust and in the technology acceptance. The second study is experimenting with the use of an AR optical see-through HMD to inform the users of the upcoming maneuvers when using rotating seats. For this study, the authors experiment different visual feedback

[187] Pampel et al., An Investigation of the Effects of Driver Age When Using Novel Navigation Systems in a Head-Up Display, in Presence, vol. 27, no. 1, pp. 32-45, March 2019.
[188] Wintersberger, et al.

to show the direction the car will turn. The conducted user study shows that when understanding the visual feedback, users have a better trust in the vehicle movements, and experience less motion sickness when not looking in the driving direction.

Call to Action

As we move forward into the future, we need to be flexible in how we approach the challenges of the AR Cloud. It will take time, in-depth research, insightful minds, and inquisitive thought leaders to unravel this new wild west.

Together as a united group of experienced designers, we can help drive and guide individuals from across the world on exciting ways to embrace the future. We must organize this effort to help promote best practices and approaches, and it can't be driven by a single company or person alone as we will need many pioneers to help us.

We highly encourage that if you are excited about helping to evolve the best in extensive experiences and interactions of the AR Cloud you reach out and get involved. The growth of this organization will help to ensure that companies listen to and adopt guidelines set forth.

We expect over time there will be a lot of breakthrough moments and that for the near term there will be much fascinating change happening rapidly. This change and active work from these groups participants will help to evolve practical use. If we have learned anything from the past, making an environment interoperable allows all the world to create re-usable experience design systems that will then become new dominant UX patterns for future generations to leverage in new and un-unanticipated ways.

As a group, we hope to push the needle forward by helping everyday people reach beyond theory and the science fiction of movies to the realm of reality. We hope this new reality will not be driven by dystopian fears but by the exciting possibilities of information, data, video, and more all available contextually when needed and when wanted.

As a group, our mindfulness of these challenges gives hope \we can lay the ground and foundations for generations to come. We would be naïve to believe this will happen overnight, or as the work of a single individual, or country. This Working Group's focus on User Experience, Accessibility, and Safety plans on tackling these challenges together.

OPEN SOURCE OPEN DATA & OPEN R&D

We look forward to having you join us in these endeavors!

Open Source, Open Data, and Open R&D

This chapter focuses on the current state of Open Source, Open Data and open Research and Development in the AR Cloud ecosystem. There were limited resources available for the research of this section, so the status overview will not be comprehensive and should rather be viewed as an initial structured probing of the topic.

Besides analyzing open source, data, and research one by one we suggest that deeper assessments could determine if an entire technological ecosystem is said to base itself on open transparent, standardized cloud compute architectures or if it should be more closed, proprietary and opaque. Could and should the ecosystem "design" itself, be an open process where architectures are shared and iterated over like open source code?

Our Inspirations were Open Edge Computing[189] and Open Compute[190].

Scope

General benefits of "open:

It is hard to overstate the positive impact open source development has had on the modern world and the modern economy.

Hundreds of years of accumulated scientific knowledge freely available to all of humanity created through the scientific method in a collaborative way has become a fundamental enabler for our technological civilization.

With the rise of machine learning and data science, data is becoming a new key enabler for new discoveries and new ways to create value.

There is a rising awareness that both R&D and data risks becoming increasingly siloed inside organizations, both in government institutions and corporations. As a response, there are many vibrant initiatives for both open research (most notable the open access science journals like arxiv.org) and data!

---

189 http://openedgecomputing.org/
190 https://www.opencompute.org/

The continuation and expansion of open source, open data and open R&D is more beneficial for more people than opaque, siloed, locked in approaches.

Benefits for the AR Cloud Ecosystem:

- **Transparency around privacy, security and safety:** Arguably the increased threat to privacy, security and safety that AR-cloud technology could introduce is a strong argument for greater transparency than earlier technology developments. In an age of "techlash" it will take more effort to build the required trust amongst stakeholders in the society for this technology to get accepted.

- **Accelerating technology development through open innovation:** The technological ecosystem of AR Cloud is both young and rather complex. One can expect a long and arduous journey through uncharted territory before it reaches the maturity needed to realize its full potential. If a culture of open innovation takes hold we will at least learn much faster from each other's mistakes and accomplishments.

- **Enabling more use-cases:** Several use-cases might only work if some types of data can be openly shared within certain contexts.

- **Open datasets for AR-cloud related machine learning:** We increasingly used machine learning to tackle a wide range of challenges related to anything from localization, mapping to machine understanding of a rich and very varied semantic, spatial, behavioural and social context around users. Better labeled datasets for training machine learning models would provide a great benefit.

- **Approved open datasets for machine learning:** There is a wide range of problems that might arise from biased or limited training data. We have seen traditional human prejudices being transferred to neural net models that trained on human generated data. In the era of spatial computing a localization method that has trained on street level data from North America might work poorly in rural India resulting in a performance bias that might leave regions of the world poorly served by solutions using such models. If there is an approval process of training datasets that aims to reduce risks of different detrimental biases and a process to amend or improve such datasets as deficiencies are discovered we could imagine a process to both approve and rate known datasets and allow developers and service providers to point towards those approvals and ratings in their “product/service declarations”.

Key Enablers

...of Open Source:

- The widespread adoption of Open Source in general
- Recognition for contribution to the economy

...of Open Data, Data Commons and Crowdsourced Open Data

- Governments have generally responded well to request to open their data
    - Government laser mapping initiatives
- Large communities of dedicated individuals who volunteer their efforts to projects like Wikipedia, Open Street Map, and creative commons content

...of Open R&D

- The Open Access Publishing in Machine Learning (arXiv)
- A movement to enable better reproducibility pushes R&D efforts to be more detailed about their methods and the data they have used to obtain their published results

Current status

Open Source:

AR technology arguably developed as early as in 1968 when Ivan Southerland demonstrated a head-mounted display named “The Sword of Damocles”. However, the ecosystem around AR Cloud is still young.

Key players like Google, Apple, Magic Leap and smaller startups are filing patents at a high rate and not necessarily in a position where their investors would be happy if they shared algorithms that might give a company a competitive edge. However, a lot of the cutting edge internal R&D that goes on inside big companies originated to a large extent in an academic institution on government-funded research grants, where key insights were discovered and shared in a far more open way.

As in AI, a lot of the top academic minds get hired on very lucrative terms by big and small tech companies where they work on advancing the technology within the confinements of the business models of each of those companies. This causes two problems: A significant brain drain from academic institutions and a natural tendency of companies to not share publicly the implementation details of breakthroughs that gives them competitive advantages. In AI, describing the performance gains characterizes the public announcement of a breakthrough, while simultaneously being quite vague on the details of how to achieve that performance, making it impossible for other experts in the field to replicate the results directly. We see a mostly of the same in AR Cloud, where both incumbents like Google describe the workings of their Visual Positioning System in rather general and vague terms, and the same is true for startups who are developing similar AR Cloud based positioning capabilities.

Open Source Initiatives

Although the general tendency in the nascent AR Cloud industry is to protect and patent their IP rather than going open source, there are some relevant open source projects and initiatives that could form the foundation of increased collaboration and more rapid open innovation going forward:

Relevant To The AR Cloud

- The Open Source Geospatial Foundation[191]
- TensorFlow (a machine learning framework)[192]
- TensorFlow Federated (training and testing on device which avoids sending private data to the cloud)[193]
- Godot Engine (a popular fully featured game engine)[194]
- OpenCV (a computer vision library)[195]
- Hudi (a big data library developed by Uber for mapping needs)[196]
- Border Go (a prototype app commissioned by the Norwegian Mapping Authority in 2017, that could display geospatial data in the AR View with centimeter-precision on a consumer grade smartphone)[197]

General

- Open Source repositories like Github and NPM
- The Linux Foundation[198]
- Open Source Initiative[199]

[191] https://osgeo.org
[192] https://www.tensorflow.org/
[193] https://github.com/tensorflow/federated
[194] https://godotengine.org
[195] https://opencv.org/
[196] https://eng.uber.com/apache-hudi/
[197] https://github.com/kartverket/bordergo
[198] https://www.linuxfoundation.org
[199] https://opensource.org/

Open Data:

Open data is a concept that is gradually gaining traction. There has been open data for centuries, particularly in the realm of science. But the modern definition came to be in 2005, introduced by the Open Knowledge Foundation who said: "Open data and content can be freely used, modified, and shared by anyone for any purpose". They defined it more elaborately in the Open Definition document currently at version 2.1[200] A notable call for action was issued by Sir Tim Berners Lee in a 2009 Ted Talk titled “The Next Web” where he called for “Raw Data Now” at the 20th anniversary of his invention of the web.

In the realm of spatial computing there are many initiatives that are noteworthy, but few have had such a wide impact as Open Street Map, which is a collaboratively created global dataset with map data from all over the world. This data powers anything from purely commercial location based apps to coordinating life saving humanitarian aid during natural disasters or wars. In fact, Tim Berners-Lee enthusiastically described in a 2010 Ted Talk titled “The Year Open Data Went Worldwide”[201] how a big volunteer effort on Open Street Map transformed fresh satellite data over earthquake stricken and poorly mapped Haiti, into much needed detailed maps of the situation on the ground. This made a huge difference in the disaster relief efforts.

However, regarding AR Cloud technology, the datasets of Open Street Map have too little detail to be of much use outside of navigation. AR Cloud technology benefits more from high resolution 3D data. In this report we refer to the two parts of the reality layer where the mostly static part comprising 3D terrain, buildings, roads and other features of the environment that rarely changes and simultaneously does not contain sensitive or private data. Therefore, the static reality layer is a great candidate to consist of open data. And luckily there is already a lot of open data of this type available. A lot of it is based on remote sensing (i.e. satellite data and aerial survey data).

Relevant to the AR Cloud

- Open Street Map (data)[202]
- Geonorge (open geospatial data from Norway)[203]
- Free global terrain datasets[204]

[200] http://opendefinition.org/od/2.1/en/
[201] https://www.ted.com/talks/tim_berners_lee_the_year_open_data_went_worldwide#t-219872
[202] https://www.openstreetmap.org/
[203] https://www.geonorge.no/en/
[204] https://gisgeography.com/free-global-dem-data-sources/

- ALOS Global Digital Surface Model[205]
- Brittish laserdata 50 cm horizontal resolutions[206]
- Høydedata (Norwegian laser based terrain data)[207]

General

- Open Knowledge Foundation[208]
- UK Open Government (data)[209]

Open R&D:

The scientific method over time naturally tends towards openness, but in the short term it's not always as open and accessible as one might imagine. Scientific journals are often prohibitively expensive for many people and demographics to access, and a lot of published scientific research is hard to reproduce due to lack of openness around methods and data and selective publishing.

We can view Open R&D as a reaction to such practices and is associated with phenomena like the "Open Access" movement that aims to make scientific literature freely available online on sites such as arXiv[210].

Also in 2017, Nature published *A Manifesto for Reproducible Science* as a response to numerous revelations of scandals of scientific malpractice. The movement strives for better reproducibility by amongst other things encouraging researchers to reveal in sufficient detail their methods and the data they used to obtain results.

Relevant to the AR Cloud

- Open Access Publishing Initiatives like ArXiv.org
- OpenAI (research)[211]

General

- Budapest Open Access Initiative[212]
- "A Manifesto for Reproducible Science", Nature [213]

---

205 https://www.eorc.jaxa.jp/ALOS/en/aw3d30/
206 https://data.gov.uk/dataset/5f6f7d5b-3f4c-4476-bfb8-cda490c9cf0e/lidar-composite-dtm-50cm
207 https://hoydedata.no/LaserInnsyn/
208 https://okfn.org/
209 https://www.opengovernment.org.uk/
210 https://arxiv.org
211 https://openai.com
212 https://www.budapestopenaccessinitiative.org/
213 https://www.nature.com/articles/s41562-016-0021

AR Cloud Related Open Organizations

Company/ organization	Category	Open Source	Open R&D	Open Data
Open Source Geospatial (OSGeo)	Non-profit	X		
Open Geospatial Consortium (OGC)	Non-profit		X	
Norwegian University of Science and Technology (NTNU)	Academic	X	X	X
Norwegian Mapping Authority	Government	X	X	X
Fantasmo.io	Commercial	X		
Mapbox	Commercial	X		
Mozilla		X		
Open Data City of Helsinki*	Commercial & Government			X

* Open Data City of Helsinki. Source: Umbra3d + City of Helsinki[214]

Direction of Evolution and Future Projects

In this report we have only just scratched the surface in this area, but we see great potential of different flavours of open for the AR Cloud industry. OARC has obtained some initial examples of relevant initiatives and organizations and we are reaching out to some of them. We have partnered with Open Source Geospatial which has been a leading organization promoting and driving a long list of key open source projects that are in wide use in the GIS sector.

OARC envisions that several projects could help drive the industry and the technology going forward. There is a case for bringing together governments, academia, NGO's, SDO, large and small industry players as well as the open source and open data community to
contribute to some key projects related to the AR Cloud:

214 https://www.umbra3d.com/clients/case-studies/city-of-helsinki/

Open Simulation and Synthesis R&D toolchain for AR Cloud Technologies

As explored in this the *Vision of Open AR Cloud* section of this report, borrowing a concept from the game engine based toolchains of the autonomous vehicle we believe we could help accelerate technological progress in this sector by creating something similar for the AR Cloud industry. One would have to adapt the toolchain to the unique and shifting technological characteristics of developing AR Cloud technology. OARC has gotten a thumbs up from the founder and lead developer of the Godot game engine they would welcome projects building such a proposed toolchain on top of their increasingly popular and capable open source game engine.

Reference Implementations for Conversion Between Geopose and Cartesian Coordinate Systems.

Given that OARC gets wider traction for our idea of creating a universal geopose standard, as described in more detail in the section *Spatial Indexing and Geopose*, there will be a need throughout many parts of the technological ecosystem to convert object poses between a geospatial coordinate-system and local ones (typically the Cartesian *x,y,z* in metric as used in most AR SDKs). We propose creating reference libraries in different programming languages as open source and will contribute towards such efforts.

Open Source Universal Spatial Browser for the Real World:

At this point in time it is too early to develop an open source project for a browser of the type described in the *Vision of Open AR Cloud* section. But given that a wider community of stakeholders at some point assembles an initial set of requirements for such a browser we definitely would like to contribute towards that end. It is hard to say before we have a better idea about the requirements for such a browser if the better route is to expand existing web browsers of if one should start from scratch to get a clean break. We are of the opinion that the concept of a browser you can run on all platforms and all devices is a superior approach than trying to offer the same capabilities as an integral part of a native OS requiring that developers to build solutions for multiple platforms to serve the whole market.

For the next year or two we expect experimentation on narrower capabilities and features be the dominant activity of those looking at real world spatial browsers, while OARC will probably focus more on fundamental standards that would be necessary regardless of browsers.

company /product	multi experien ce	collabor ation	Content Sources	platforms	Model
YouAR	Yes	collabora tive interacti on	YouAR proprietary interoperable contents	android, IOS, Windows Mixed Reality	Proprietary development tools
Wikitude browser	no	shared content	Wikitude proprietary solution	android	Free android application
Vuforia View	No	shared content	Vuforia Studio	android, ios, AR glasses	Free, but that is not the case of contents creation tools
ContextGrid's HoloGrid Browser	Yes	No	Can link multiple data sources for content	IOS Only for now, Android and Hololens coming soon	Free browser, createor pays per content placement.
WebXR	Have to be impleme nted by designer s	Have to be impleme nted by designer s	Alll standard content types from all online sources	the web (extending the webplatform to include spatial browsing)	Web development standard

Early spatial browsers (mostly not open-source)

Open Dataset on the "Static Reality-Layer".

We are very encouraged by The growing set of open data on terrain, buildings and the like encourages us.There are whole cities, like Helsinki, mapped and shared openly. The next natural step would be to try to bring such datasets together in a global repository and converting the data into standardized formats and making the data easily available via web services. Eventually such a project could end up with a comprehensive 3D data set of the whole world that AR Cloud service developers could use as base-data. This dataset could lay the base foundation for a static reality-layer as we have described in *Vision of Open AR Cloud*.

Projects like Open Street Map have shown the way and future projects could model a similar approach. But this will bring in a challenge related to the sheer amount of data and the requirement of low latency access to that data. Edge computing is in an early phase, the same goes for Distributed Ledger based decentralized data storage, but we envision those types of technologies could be part of both creating the dataset, and distributing it cost efficiently to where they will be of most frequently used and valuable. As those types of technology matures the different options for the architecture and handling of the data will become clearer.

AR - Open Source Resources (not intentionally aimed at AR Cloud)

- AR.js[215]
- ARToolKit[216]
- Mark Billinghurst (Prof. in HCI at University of South Australia, + venture partner Superventures) list of where AR-research is being done[217]
- The Linked Open Data Cloud (LOD)[218]
- AR Research related to LOD[219]:

[215] https://github.com/jeromeetienne/AR.js/blob/master/README.md
[216] https://github.com/artoolkit
[217] https://medium.com/@marknb00/where-in-the-world-is-ar-vr-research-happening-ddebbdc6436b
[218] https://lod-cloud.net
[219] https://www.researchgate.net/publication/322337536_Linked_Open_Data_as_universal_markers_for_Mobile_Augmented_Reality_Applications_in_Cultural_heritage

Distributed Ledger Technology and Blockchain

Introduction

Perhaps “AI” and “Machine Learning” are the only two terms that might better evoke excited hyperventilation in Silicon Valley, other than “Augmented Reality” and “blockchain”. But you cannot combine AI and machine learning to describe a breathtaking new crossroads of technologies because those two terms are related, if not directly synonymous. However, the crossroads of augmented reality and blockchain is, in fact, a real destination, albeit a destination that no company or project is successfully navigating (yet).

As explained elsewhere in this report, the AR Cloud is a technology stack that generates a media channel that exists as a digital content layer(s) indexed to the real world. Rather than tuning to, say, YouTube for a music video, or MSNBC for news, one might - using a mobile phone or HUD AR system - “tune in” to their neighborhood street art mural by walking to it or a magazine page by opening to it, for digital stories or experiences that revolve around that real world location or thing. But the UX of constantly (and consciously) switching between layers as you walk through life is fatiguing and suboptimal. Instead, we will need to curate the AR layered content in some fashion, either as knowledgeable tastemakers, or with AI agent systems that have been trained on our preferences—personal agents that have the capability to sift through the digital layers and select the ones it thinks will resonate with you.

Nonetheless, whether it’s content presented by automated agent or by tastemaker, AR users will not want to constantly be presented with manual choices on all the content available to them and will not want to engage in a transaction to download or experience these potentially microexperiences or story vignettes. The content stitched around us via the AR Cloud demands a transactional backbone that is fully digital and automated—that transactional backbone could be blockchain technology.

Scope

Background

Blockchain technologies are a formative innovation for the Internet and AR Cloud landscapes. Like other foundational transfer protocols that send website or email data packets successfully across the entire Internet, blockchain technologies enable the transfer and tracking of ownership of specific packets of data across the Internet. For example, Bitcoin and Ethereum each enable the permanent transfer of ownership of different data packet types - Bitcoin enables the trustless exchange of ownership of financial instruments, and Ethereum enables access to or the transfer of ownership of distributed computational power. Unique to blockchain technologies, the ownership of each individual packet of data is added to a global distributed ledger (i.e. a publicly or privately verified database) that assigns ownership from one particular account (a transactional address) to another. Following is the definition of Digital Ledger Technology according to Wikipedia:

"A distributed ledger (also called a shared ledger or distributed ledger technology or DLT) is a consensus of replicated, shared, and synchronized digital data geographically spread across multiple sites, countries, or institutions. There is no central administrator or centralized data storage. A peer-to-peer network is required as well as consensus algorithms to ensure replication across nodes is undertaken.

A blockchain is a growing list of records, called blocks, which are linked using cryptography. Each block contains a cryptographic hash of the previous block, a timestamp, and transaction data."

This ledgerization of the ownership of data prevents the typical digital "double-spend" problem - where it had been impossible to guarantee that a specific data file was permanently and exclusively transferred from one computer to another, without retaining a copy of that file.

Alignment with OARC

The concept of these distributed ledgers is very much in alignment with the Open AR Cloud initiative the ledgers are distributed across many (sometimes thousands) computers, ensuring that no single entity controls the data embedded in the ledger. The record of ownership is democratized across each system, and typically, the larger the system of ledger computers, the more secure and robust it is. Control over the system is embedded in the blockchain code architecture, which is maintained and updated by a similarly distributed group of contributing developers. Sometimes that group is centralized itself under a governing board or council, like Ethereum. But sometimes that group of contributing developers runs a spectrum of variables, including whether it is: entirely volunteer, contribution or merit-based, centralized-control or majority vote based—resulting in sometimes contentious administration of the blockchain code.

Decentralized vs. Centralized

The decentralized nature of DLT is a compelling solution to the conflict caused by the centralization (monopoly or pseudo-monopoly) of transactional economies.

Top Benefits of Blockchain[220]

1. Greater Transparency
2. Enhanced Security
3. Improved Traceability
4. Increased Efficiency and Speed
5. Reduced Costs
6. Democratized Governance and Control

Types

Bitcoin was the first blockchain technology deployment, and it uses a coding language called Script which provides for some very limited data types and functions via the Bitcoin blockchain. Each blockchain technology type (there are over 3000 currently) is optimized for a specific kind of data type or use case, and this data transfer is often embodied in an access token. These access tokens serve many functions in the system: 1) they create technological gateways through which participants must use the token to enter, ensuring that the user has the correct data type for the network; 2) tokens create scarcity of access to the network, ensuring that users have an invested value in accessing the network with their data; 3) the scarcity function drives inherent value in possessing the token; if you possess an access token to a network that is congested and in heavy demand, the market price for your token goes up (assuming there is a ready market for it).

I. For Bitcoin, the public exchange symbol for its token is BTC, and for Ethereum, the symbol is ETH. And there are some blockchain initiatives designed to bridge across different token systems—effectively allowing Bitcoin data packets to be interpreted on the Ethereum blockchain, or vice versa—and across as many of the 3000-plus technology stacks as can be implemented.

Non-Alignment

Because of this fractured landscape, and because the financial incentives embedded into each blockchain system can be very large, this space can at times run at odds with the OARC goal of open access—some tokenized blockchain systems will want siloed winner-take-most control of that data type. To the extent that any one blockchain gains early and prominent control over a data transaction type that is foundational to the AR Cloud, OARC needs to take a "bridging protocol" approach to working with that embedded token system—using leverage to pressure access to the token system in a manner that allows for interoperability with other

[220] https://www.ibm.com/blogs/blockchain/2018/02/top-five-blockchain-benefits-transforming-your-industry/

transactional tokens in the OARC system - while at the same time fostering development of competing open protocols that may one day supplant the embedded control token.

Key Enablers

- Mobile GPU Efficiencies and Edge Computing
- 5G Latency and Bandwidth
- Economic Incentives
- Reduced IRL Fees by Cutting Out Payment Processors
- Mobile Wallets & dApps
- Data Marketplace with Ownership and Trust

Key Challenges

- Wallet UX (e.g. Ethereum Gas)
- Interoperability
- Performance (latency, functionalities supported,)
- Scalability (transactions per second, global infrastructure, energy requirements)
- Processing Cost Externalities (heat, cost, environment)
- Geopolitical Forces (mining centralization in a given country)
- Fractured Landscape
- Governance (within blockchain forks)

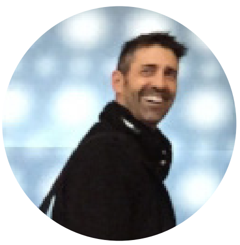
Current status

Critiques
Banning
Environmental concerns
Hacked exchanges
51% Attack
What we're not going to embrace
Cryptography
Blockchain Use Cases - Collin WP

Key to Columns in Vendor Spreadsheet

https://docs.google.com/spreadsheets/d/1huzWYsIWeXppw1sjF064ViFs1wx8k2hAhNKI4aIPKpM/edit?usp=sharing

dave lorenzini:
Dave Lorenzini is a spatial computing expert, author, speaker, awe/augmentedreality.org board member, and co-founder of openarcloud.org. He has helped create, build and lead xR, media and mapping companies that have changed the way billions of people look at the world. Successes include keyhole (google earth), space imaging (maxar) and ARc, where he led xR product, platform and strategy work for clients like googleX & facebook bld. 8. Dave lives in scottsdale and south lake tahoe, works in the bay and la, and can be found online everywhere as: @davelorenzini.



Chris Nunes:
Chris Nunes is the Head of Studio at HEAVY (www.heavy.io), an award-winning AR creative lab with partner BC Biermann. The HEAVY team specializes in interactive, large format, public art installations, but the team also consults with enterprise and retail clients on prototypes, creative campaigns, and AR-first experiences, while also developing their own portfolio of AR apps. HEAVY's clients include global household technology brands like Google and Tesla, public & private art committees like Miami Art in Public Places and Atlanta's WonderRoot, cities & governments like the City of Moscow and the European Union Commission on Civil Rights, schools & institutions like the University of Geneva and Voices of the Children, and real estate developers, architects & fine artists. Inside HEAVY, Chris runs all client and product decisions. He was formerly an entertainment attorney in Los Angeles, past chair of the IEEE Augmented Reality Industry Connections initiative, and chair emeritus of the SHC Student Launch Initiative, an entrepreneurial incubator at Chris's high school alma mater in San Francisco. He now resides in Boulder CO with his family and travels the world with HEAVY helping communities rediscover the wonder of their shared spaces through technology art.

Rolf Kruse
Rolf Kruse is Professor for Digital Media and Design at the University of Applied Sciences in Erfurt, Germany since 2012. Originally educated as architect he came in contact with VR in the early 1990s at the Fraunhofer Institute in Darmstadt helping to boot the first VR-Center. Since then he has a long experience in the ideation and implementation of interactive spatial media with innovative technologies. Be it in the area of city planning (at Art+Com, Berlin, 1994-1997), mixed reality serious games for tradeshows, mediatecture installations (with Invirt GmbH, 1997ff.) or the "Cybernarium Edutainment Center" (a Fraunhofer Spin-Off in Darmstadt 2002-2005) – always using newest media technologies to transfer complex knowledge to empower people.
His research focuses on Virtual-/Augmented-/Mixed-Reality as human computer interfaces for intuitive interaction with dynamic spatial information that are seamlessly integrated into our environment. Since many years he is active part of the academic german community (GI VR/AR + E-Learning) as conference chair and reviewer.

Jason Fox
Jason Fox is a spatial computing engineer in Microsoft's Commercial Software Engineering organization, where he works with enterprises to build robust spatial computing systems with technologies such as HoloLens, mobile AR (iOS/Android) and accompanying cloud services.

In the past, Jason has designed and built 3D visualization systems, built video game systems, non-destructive testing systems, spent time as a technical evangelist and has developed real-world solutions to help solve humanitarian needs.

Jason is currently a contributing member to the Open AR Cloud organization where he chairs the reality modeling and mapping working group.

Christine Perey
Christine Perey is an industry analyst and active leader of new technology industry initiatives. She was an early evangelist for use of streaming media and videoconferencing for 15 years until, in 2006, she began to study mobile Augmented Reality to increase operational efficiencies in business and enrich the lives of people. Her company provides services and programs for executives building and acquiring AR-enabling technologies as well as to the largest enterprise AR platform customers.

Christine has started and led many communities of interest. She is an outspoken advocate for interoperable Augmented Reality content and experiences and, from 2009-2016, led a grassroots community dedicated to this purpose. Christine serves on numerous standards working groups and is vice-chair of IEEE P1589. She is the founder, and from 2013 to 2016 served as the founding executive director of, the AR for Enterprise Alliance (AREA), the only global member-based organization accelerating AR adoption in enterprise.

When not traveling, Christine lives and works in Montreux, Switzerland.
https://www.linkedin.com/in/christineperey/

Tony Hodgson
Tony Hodgson is Founder/CEO of Brainwaive LLC, a global advisory to enterprise and tech startup clients on AR/VR, emerging technology, and cyber security. He's a multi-degreed
engineer who supported NASA and commercial space programs, then became a global telecom solutions lead, serial tech entrepreneur, and executive activist for development of open, interoperable and secure AR/VR solutions. In 2017, the Brainwaive team led development of a seminal cyber security framework and test protocol for enterprise AR applications for the AR for Enterprise Alliance (AREA). Brainwaive currently leads the Security Committee of the AREA, and the Privacy & Security Working Group of the Open AR Cloud initiative. Tony helps companies thrive where emerging tech and business opportunities converge.

James Jackson
James Jackson is a Principal Member of Technical Staff at AT&T Labs in Austin, Texas. He primarily focuses on distributed systems design of cloud-based mobility messaging platforms and also has a deep interest in services innovation and rapid software prototyping.

More recently, his interests extend to transformative applications of computer vision and machine learning including self-driving cars and augmented reality. James strives to promote open and interoperable technologies with a consideration for social good applications.

James received a BSEE from the University of Texas at Austin and an MSCS from Georgia Tech. He is a contributing member of Open AR Cloud and co-chairs the reality modeling and mapping working group.

Marco Tillmann
Marco Tillmann is a product manager for HERE Technologies, a global leader in mapping and location platform services. Marco specializes in augmented and virtual realities as well as 3D cities. He is responsible for driving-innovation and life-cycle management for 3D visualization and immersive technology products.

Prior to HERE Technologies, Marco received his degree in computer art and 3D character animation at the Savannah College of Art and Design, Savannah, GA and led development of 3D authoring products at MAXON Computer GmbH in Friedrichsdorf, Germany, for the VFX industry, as well as supporting developer relations at Apple Ltd in Uxbridge, UK and product management and sales support at Integrated Computing Engines Inc, in Waltham, MA, USA.

Marco also serves as Co-Chair of the Spatial Indexing and Geopose Working Group of the Open AR Cloud.

Scott Simmons
Mr. Simmons is OGC's Chief Operations Officer. In this role, he provides oversight and direction to the Consortium's technical and program operations and deliverables. Scott also continues to lead the OGC Standards Program, where he ensures that standards progress through the organization's consensus process to approval and publication.

Preceding his time as a member of OGC staff, Scott was an active member of OGC, promoting best practices in 3D Information Management (3DIM) as chair of the OGC 3DIM Domain Working Group and chairing or participating in numerous OGC infrastructure, mobility, and web services working groups. Scott was a contributor to the OGC Augmented Reality Markup Language (ARML 2.0) standard. His OGC-related research has focused on data lifecycle management, integration, and dissemination.

Preston McCauley
Preston is a co-founder of the UX program at Southern Methodist University (GO) in Dallas, Texas where he has taught Practical user experience strategy & process methodology for five years. While teaching at SMU GO, he has collaborated to create new course curriculum for a new elective course, UX Principles of AR & VR.
He believes in sharing knowledge in as many ways as possible with the design and development community in person and online. He is frequently visible speaking, teaching, and educating people on the latest in bleeding edge technology, design, prototyping, and developing skills in the fields of immersive and next level system design. To further this objective, he has co-founded the Dallas AR / VR Design & Development group, which currently has over 900 members.

For the last several years, he has been researching, participating, and formulating new theories on interaction design to help usher in the next era of AR / VR design for UX professionals.
His other accomplishments across his career include:

20 Years As A UX Professional
Developing His Own UX Process Strategy & Methodology Approaches
Co-Hosting Podcast Augmenting Reality
AR / VR / UX mentor for SXSW in Austin, Texas for two years
Professional speaker at over twenty industry events
Creating online courses to acclimate others to the rise of AR / VR
Joining the OARC UX Working Group to help define the next global state of dimensional design - the AR Cloud.

He is currently working as Principal UX AR / VR Innovation & Design at Tonic3 in Lewisville, Texas Owner Clear Sight Designs,LLC - You can find him on twitter @uidesignguide

Michael Vogt

Michael Vogt started his programming career at Apple AICS in Germany, specializing in QuickTime and video streaming. After publishing a book about QuickTime, worked on video recording and streaming platforms. He has since worked as a freelancer on many projects for the consulting, telecommunication and banking industry.

He currently concentrates on applications of AR for tourism and is member of the working group reality modelling, mapping and spatial data creation.

Alina Kadlubsky
Alina Kadlubsky is an Art Director, Digital Designer Web XR Developer.
A Young Creative, with a passion for immersive Tech, Spatial Web & AR.
She is Co-Founder of Matrixcore, an initiative that partnered with AWE for an interactive Virtual Conference experience 2019.
Joined the Open AR Cloud Family November 2018, involved in the rebranding and Inaugural Working Group Kickoff process. She is also Chair of the User Experience Accessibility and Safety Working Group.

John Daly
John Daly is the Program Manager for Immersive Technologies and NAVAIR Programs at IPKeys Technologies in TInton Falls, New Jersey. John's focus is on translating training into engaging experiences for education products serving military, medical, and federal agencies.

John is the project lead for the US Military's immersive trainer, I-GAME, which has been accredited and validated as pre-deployment training at 13 US installations worldwide instructing warfighters on Counter-IED techniques, tactics, and procedures.

John is a graduate from Fairfield University, a Google Street View Trusted photographer, and Certified Scrummaster with the Scrum Alliance. He is a contributing council member to Open AR Cloud Working Group 8 excited about the potential of well designed and accessible AR experiences.

Suzan Oslin

Suzan Oslin is an eXperience aRchitect and product strategist specializing in immersive technologies. She evangelizes human centered design in the VR / AR / MR communities locally as a co-organizer of the LA UX Meetup and AWE Nites LA, and contributes globally through her work with the Open AR Cloud. Her career spans over 20 years in digital media with origins in visual effects animation at Disney and Warner Brother's Feature Animation and digital product design for Technicolor and Deluxe Media Services. Suzan focuses on the impact that technology has on the evolution of humankind and believes through interdependent collaboration and mindful use of emergent technology we can create a better tomorrow for all.

Julien Casarin

Julien Casarin is the Head of eXtended Reality at Gfi Innovation in Paris. His work is mainly focused on improving the support of cooperative work with immersive technologies. Julien also is the coordinator of the UMI3D Consortium creating an Open source protocol designed for cross-device collaboration in AR/VR/MR. Finally, he is completing a PhD thesis on collaborative VR at the University of Strasbourg.