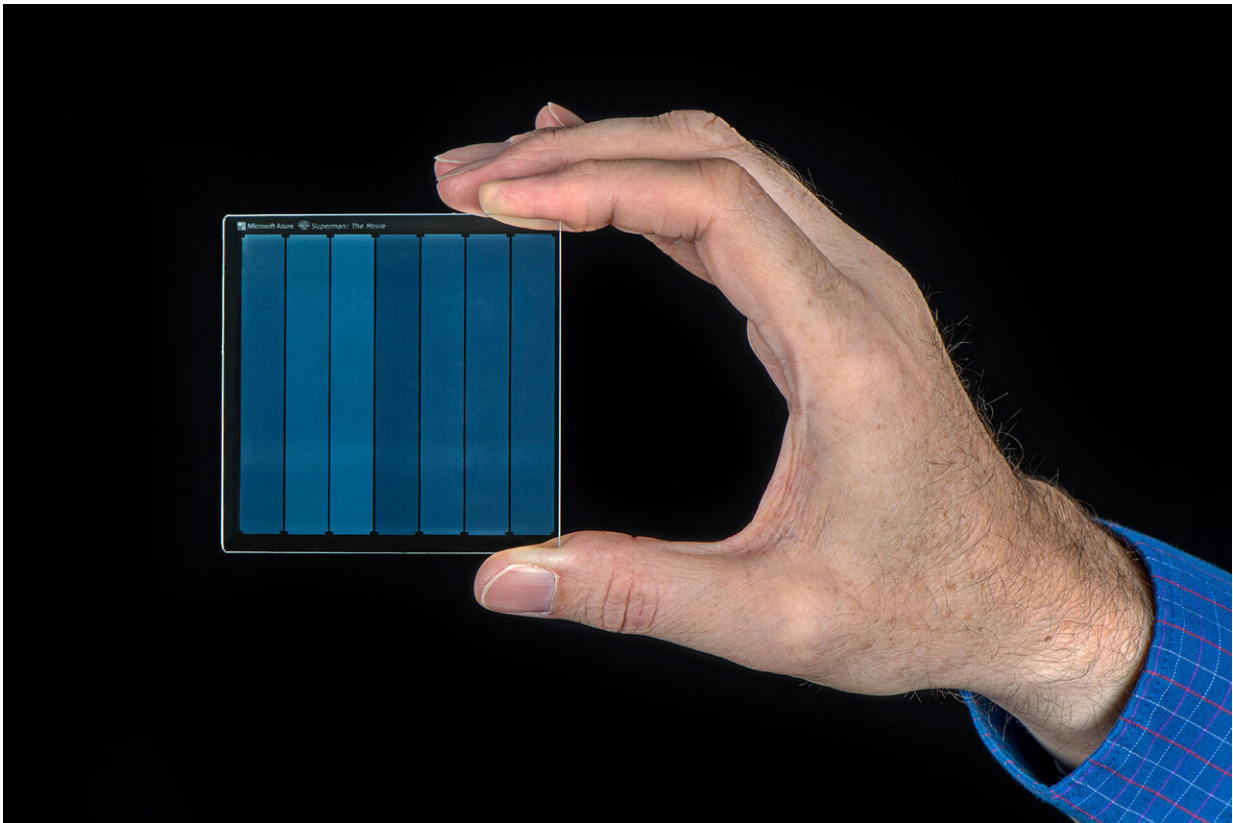# Project Silica proof of concept stores Warner Bros. 'Superman' movie on quartz glass

November 4 2019, by Jennifer Langston

In a proof of concept test for Project Silica, Microsoft and Warner Bros. have successfully stored and read back the classic 1978 "Superman" movie on a 7.5 cm x 7.5 cm x 2 mm piece of durable silica glass — roughly the size of a drink coaster. The glass contains 75.6 GB of data plus error redundancy codes. Credit: Jonathan Banks for Microsoft.

Microsoft and Warner Bros. have collaborated to successfully store and retrieve the entire 1978 iconic "Superman" movie on a piece of glass roughly the size of a drink coaster, 75 by 75 by 2 millimeters thick.

It was the first proof of concept test for Project Silica, a Microsoft Research project that uses recent discoveries in ultrafast laser optics and artificial intelligence to store data in quartz glass. A laser encodes data in glass by creating layers of three-dimensional nanoscale gratings and deformations at various depths and angles. Machine learning algorithms read the data back by decoding images and patterns that are created as polarized light shines through the glass.

The hard silica glass can withstand being boiled in hot water, baked in an oven, microwaved, flooded, scoured, demagnetized and other environmental threats that can destroy priceless historic archives or cultural treasures if things go wrong.

It represents an investment by Microsoft Azure to develop storage technologies built specifically for cloud computing patterns, rather than relying on storage media designed to work in computers or other scenarios. It's just one of many ways Azure relies on Microsoft's Research expertise to solve both near- and long-term challenges—from Project Natick's underwater data center tests to Project Brainwave's FPGA processing power and the emerging Optics for the Cloud research.

"Storing the whole 'Superman' movie in glass and being able to read it out successfully is a major milestone," said Mark Russinovich, Azure's chief technology officer. "I'm not saying all of the questions have been fully answered, but it looks like we're now in a phase where we're working on refinement and experimentation, rather asking the question 'can we do it?'"

Warner Bros., which approached Microsoft after learning of the research, is always on the hunt for new technologies to safeguard its vast asset library: historic treasures like "Casablanca," 1940s radio shows, animated shorts, digitally shot theatrical films, television sitcoms, dailies from film sets. For years, they had searched for a storage technology that could last hundreds of years, withstand floods or solar flares and that doesn't require being kept at a certain temperature or need constant refreshing.

"That had always been our beacon of hope for what we believed would be possible one day, so when we learned that Microsoft had developed this glass-based technology, we wanted to prove it out," said Warner Bros. Chief Technology Officer Vicky Colf.

Driving down costs of long-term storage

Most people think of "the cloud" as a way to store everything from thousands of family photos to millions of emails without taking up any space on your phone or computer. But all that information is being physically stored on hardware in a remote location, allowing you to access it from multiple devices.

The amount of data humankind is now looking to store—from medical records to funny cat videos to images taken by spacecraft—is exploding at the same time that capacity of existing storage technologies is flattening.
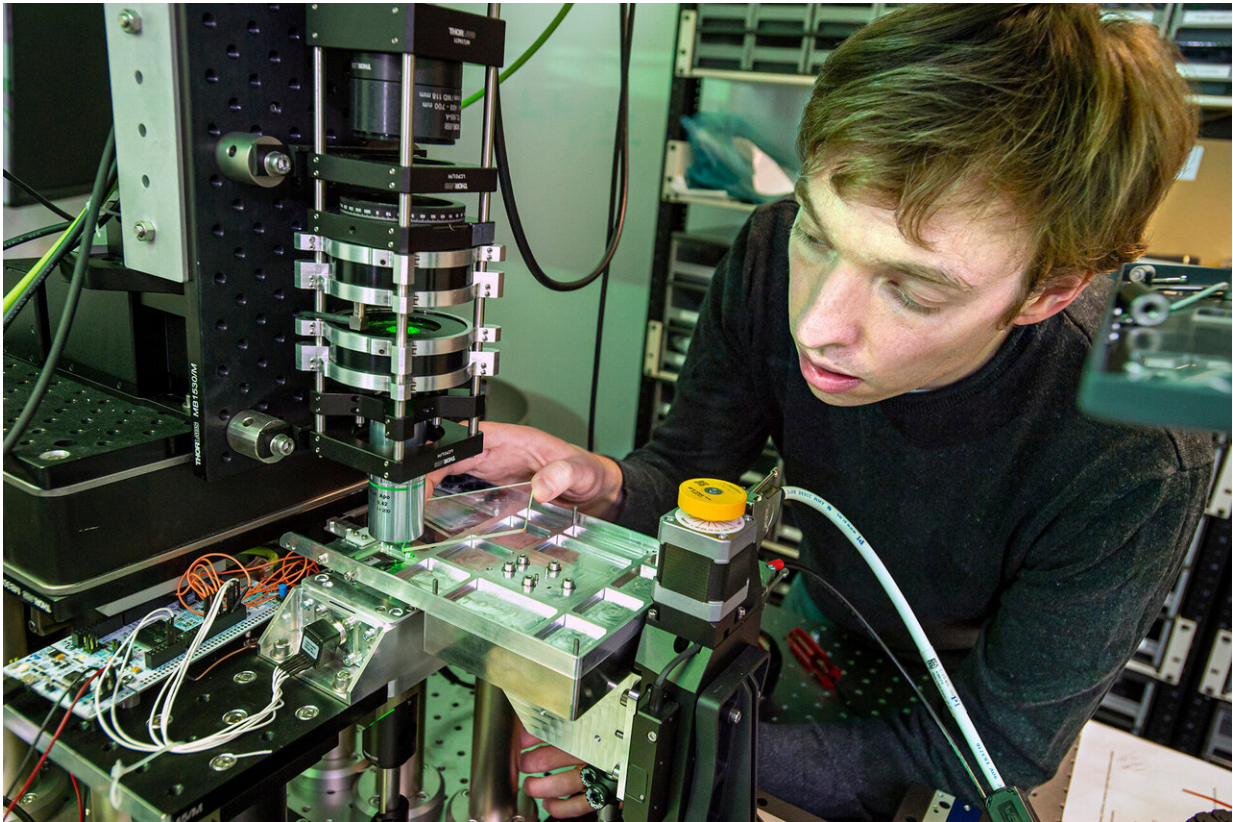
Microsoft senior optical scientist James Clegg loads a piece of glass into a system that uses optics and artificial intelligence to retrieve and read data stored on glass. Credit: Jonathan Banks for Microsoft.

Long-term storage costs are driven up by the need to repeatedly transfer data onto newer media before the information is lost. Hard disk drives can wear out after three to five years. Magnetic tape may only last five to seven. File formats become obsolete, and upgrades are expensive. In its own digital archives, for instance, Warner Bros. proactively migrates content every three years to stay ahead of degradation issues.

Glass storage has the potential to become a lower-cost option because you only write the data onto the glass once. Femtosecond lasers—ones that emit ultrashort optical pulses and that are commonly used in LASIK

surgery—permanently change the structure of the glass, so the data can be preserved for centuries.

Quartz glass also doesn't need energy-intensive air conditioning to keep material at a constant temperature or systems that remove moisture from the air—both of which could lower the environmental footprint of large-scale data storage.

"We are not trying to build things that you put in your house or play movies from. We are building storage that operates at the cloud scale," said Ant Rowstron, partner deputy lab director of Microsoft Research Cambridge in the United Kingdom, which collaborated with University of Southampton to develop Project Silica.

"One big thing we wanted to eliminate is this expensive cycle of moving and rewriting data to the next generation. We really want something you can put on the shelf for 50 or 100 or 1,000 years and forget about until you need it," Rowstron said.

Project Silica aims to store what's known as "cold" data—archival data that may have tremendous value or that companies are required to maintain—but that doesn't need to be frequently accessed. That might include medical data that has to be kept for a patient's entire life, financial regulation data, legal contracts, geologic information that pertains to energy exploration and building plans that cities need to hold onto.

Warner Bros. was keenly interested in helping Microsoft test solutions that might alleviate the costs and inefficiencies associated with storing data over these long time horizons, Colf said.

"With the largest content library in the media and entertainment industry by many measures, our challenges are unique in their scale, but they are

certainly not unique in terms of the problem we are trying to solve," she said.

Turning digital data into physical artifacts

With a nearly 100-year history in film and television, Warner Bros. owns one of the world's deepest and most significant entertainment libraries. Re-releasing older films in new formats or for new audiences is an important part of the business. It's also a tremendous cultural responsibility to preserve some of the world's most beloved stories in perpetuity, Colf said.

"Imagine if a title like the 'Wizard of Oz' or a show like 'Friends' wasn't available for generation after generation to enjoy and see and understand," she said. "We think that's unimaginable, and that's why we take the job of preserving and archiving our content extremely seriously."

The company has redundancy plans in place to handle multiple worst-case scenarios: an earthquake or hurricane that strikes one of the coasts, a fire where the suppression systems don't kick in or a climate control failure that allows moisture to build up and ruin film stock.

The goal is to have three archival copies of each asset stored in different locations around the world: two separate digitized copies, along with the original physical copy on whatever medium a film or television episode or animated cartoon was created.

Fortunately, original film negatives will last for centuries if stored in the right conditions. But for some older television shows—think episodes of "Alice" shot in the 1970s—the original physical copy has a limited shelf life that requires migration to newer formats. And for today's films and television shows that are shot digitally, the archival-quality third copy

has a very short migration cycle of three to five years, which is challenging to manage.

"Let's say a TV show is pushing directly into our digital archives; there's nothing physical," said Steven Anastasi, Warner Bros. vice president for global media archives and preservation services. "The digital file is going in but I don't have something I can put in a vault or in a salt mine or anything physical coming into the building."

Warner Bros. is potentially looking at Project Silica to create a permanent physical asset to store important digital content and provide durable backup copies. Right now, for theatrical releases that are shot digitally, the company creates an archival third copy by converting it back to analog film. It splits the final footage into three color components —cyan, magenta and yellow—and transfers each onto black-and-white film negatives that won't fade like color film.

Those negatives are put into a cold storage archive. In these highly managed vaults, temperature and humidity are tightly controlled, and air sniffers look for signs of chemical decomposition that could signal problems. If they need the film back, they must reverse those complicated steps.

That process is expensive, and there are only a handful of film labs left in the world that can do it. And the process is not optimal from a qualitative point of view, said Brad Collar, Warner Bros. senior vice president of global archives and media engineering.

"When we shoot something digitally—with zeros and ones representing the pixels on the screen ¬— and print that to an analog medium called film, you destroy the original pixel values. And, sure, it looks pretty good, but it's not reversible," Collar said.

"If we can take the digital representation of those pixels and put it on a medium like silica and read it back off exactly as it was when it came out of the camera, we've done our preservation job to the very best of our ability. That's what I love about this," he said.

Brad Collar, Warner Bros. senior vice president of global archives and media engineering, left, and Vicky Colf, Warner Bros. chief technology officer, demonstrate the contrast between storing films on 22 reels of film versus a coaster-sized piece of glass. Credit: John Brecher for Microsoft.

It's not economical to create archival film negatives for every digitally shot television episode in the Warner Bros. library. The company hopes Project Silica might prove to be a cheaper, higher quality alternative to

create physical archives of digital content.

There's much more work ahead to reach that scale—Microsoft researchers would need to significantly increase the speed at which data can be written and read, as well as its density. Warner Bros. envisions its own infrastructure to read data from the glass archives. But both partners see promise in how far they've come.

"If Project Silica's storage solution proves to be as cost-effective and as scalable as it could be—and we all recognize it's still early days—this is something we'd love to see adopted by other studios and our peers and other industries," Colf said.

"If it works for us, we firmly believe that this will be a benefit to anyone who wants to preserve and archive content," she said.

Designing storage for the cloud

It's impossible to know how much information has been lost because no one realized its value at the time—from silent movies that no one envisioned would ever be seen outside a theater to historic data that modern analytic tools and AI might glean new insights from.

One goal of Microsoft's next generation storage research, which includes parallel efforts to store data in DNA, is to develop solutions that are cheap enough and effortless enough that you don't really have to make a choice about whether to store your data, the company says.

Microsoft researchers spent years trying to get there with technologies currently used in data centers. But the size, shape and constraints of things like spooling tape and spinning disks—all of which were invented for other purposes long before the cloud existed—simply couldn't get them the gains they wanted.

"Eventually, we just thought 'can we build something from the ground up for the cloud that doesn't need to do anything else?'" Rowstron said.

They launched a collaboration with the University of Southampton Optoelectronic Research Centre, where researchers originally demonstrated how to store data in glass with femtosecond lasers. With investment from Azure, Microsoft's Cambridge, UK, lab built an interdisciplinary team of physicists, optics experts, electrical engineers and researchers with storage backgrounds to push the technology further.
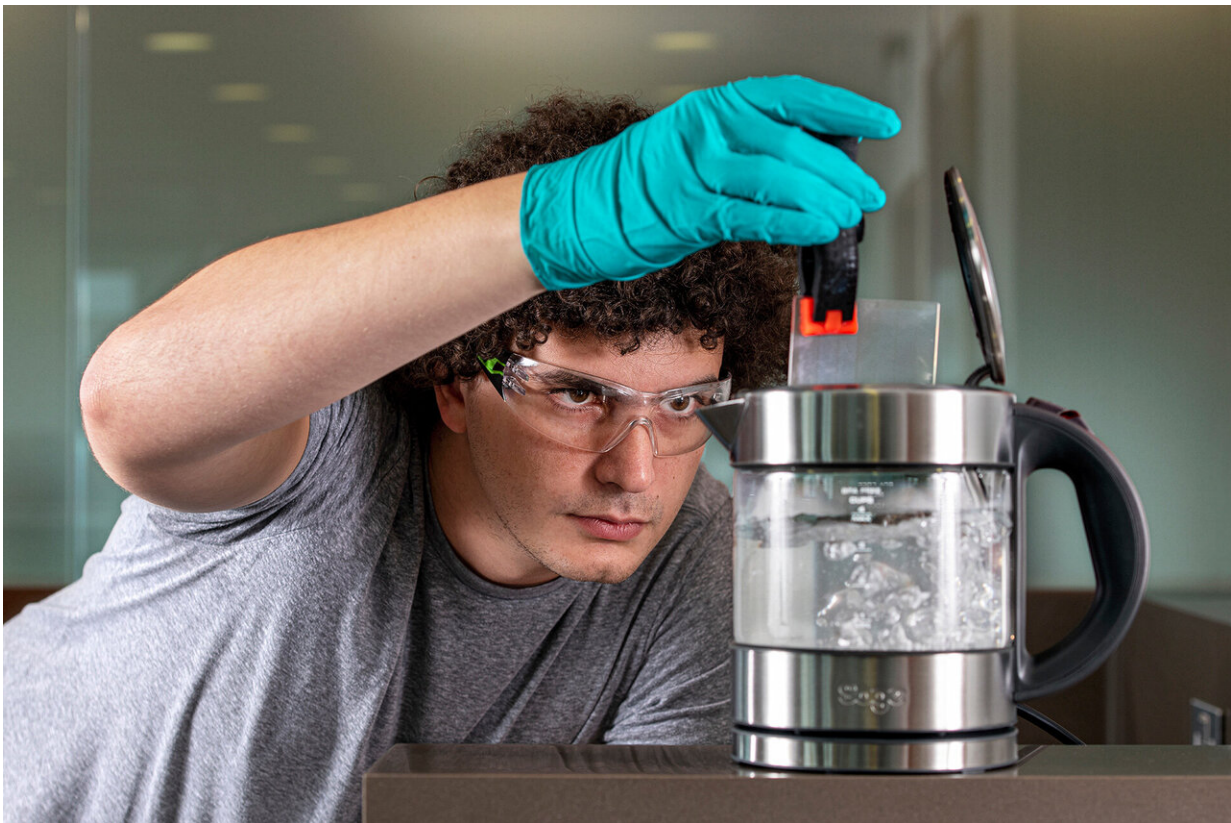
Microsoft Project Silica researcher Youssef Assaf drops a square of silica glass in a kettle of boiling water to demonstrate its durability. The team has baked, boiled, microwaved, demagnetized and scoured similar pieces of glass with steel wool — with no loss to the data stored inside. Credit: Jonathan Banks for Microsoft.

Since then, the Microsoft Research team has achieved dramatic advances in speed and precision. They've also worked closely with their Azure counterparts to design Project Silica with the day-to-day challenges and requirements of commercial cloud storage in mind.

"Getting all their input and thinking into the project from Day 1 means we're going to generate something at the end that's really usable for them," Rowstron said of the relationship with the Azure product team.

Project Silica's infrared lasers encode data in "voxels," the three-dimensional equivalent of the pixels that make up a flat image. Unlike other optical storage media that write data on the surface of something, Project Silica stores data within the glass itself. A 2-mm-thick piece of glass, for instance, can contain more than 100 layers of voxels.

Data is encoded in each voxel by changing the strength and orientation of intense laser pulses that physically deform the glass. It's somewhat like creating upside down icebergs at a nanoscale level, with different depths and sizes and grooves that make them unique.

To read the data back, machine learning algorithms decode the patterns created when polarized light shines through the glass. Unlike tape storage—which takes time to spool to get to the place you want to read back—the algorithms can quickly zero in on any point within the glass square, potentially reducing lag time to retrieve information.

"If you're old enough to remember rewinding and forwarding songs on cassette tapes, it can take a while to get to the part you want," said Richard Black, Microsoft principal research software engineer. "By contrast, it's very rapid to read back from glass because you can move simultaneously within the x or y or z axis."

Unlike fragile wine glasses or light bulbs, the squares of quartz glass used for data storage are surprisingly hard to destroy. Early on, the research team tried baking one in an oven at 500 degrees, microwaving, boiling it, scouring it with steel wool. And when they read the data back, it was all still there.

That made total sense to the Warner Bros. archivists, who years ago discovered boxes of Superman radio serials recorded in the 1940s on record-sized pieces of glass.

"We actually found players that we could play these things back on, and they were just as good because they were stored on glass. And we were able to digitize and save those wonderful pieces of content," Collar said.

"So now one of our oldest assets in our vault is glass and one of the newest technologies in our vault is glass. And they're both Superman. So we really have come full circle," he said.

Provided by Microsoft