Hi reddit!

Our team has completed a digital analysis of the extremely fragile En-Gedi scroll — the oldest Pentateuchal scroll in Hebrew outside of the Dead Sea Scrolls — revealing the ink-based writing hidden on its untouchable, disintegrating sheets, without ever opening it. While prior research has successfully identified text within ancient artifacts, the En-Gedi manuscript represents the first severely damaged, animal skin-based scroll to be virtually unrolled and non-invasively read line by line.

The series of digitization techniques we employed demonstrates that it is possible to “see” ink-based text within an extremely fragile scroll while avoiding the need for physical handling. The traditional approach of unrolling a scroll and pressing it flat in order to duplicate text is not an option for splintering manuscripts like the En-Gedi scroll, which has been burned and crushed into lumps of charcoal.

We began by performing a volumetric scan of the scroll using X-ray microtomography, followed by segmentation, which digitally creates a “page” containing the writing. We pieced together over 100 such scanned segments of the scroll by hand. Further manipulation of the digitized scroll involved using texturing and flattening techniques, and finally, virtual unwrapping to unveil the text written on its pages.

At last, we were able to “see” the text on five complete wraps of the En-Gedi scroll, and the resulting image is one of two distinct columns of Hebrew writing that contain legible and countable lines, words, letters, and spacing. Further analysis revealed the scroll’s writings to be the book of Leviticus, which makes it the earliest copy of a Pentateuchal book ever found in a synagogue’s Holy Ark. This virtual unlocking of the En-Gedi scroll paves the way for further scholarly analysis of this and other text buried in delicate, damaged materials.

Our research was published yesterday in Science Advances, the open-access journal of the American Association for the Advancement of Science. Here is our article: “From damage to discovery via virtual unwrapping: Reading the scroll from En-Gedi” Brent Seales, professor and chairman in the department of computer science at the University of Kentucky Michael Segal, the Otsuki Professor of Biblical Studies and head of the School of Philosophy and Religions at Hebrew University of Jerusalem Seth Parker is the Project Manager on the Scrolls Project, directly overseeing software development by the team’s 8 student developers. He’s also a big fan of Whit Stillman and Ross McElwee.

We’ll be back at 11 am EST (8 am PST, 4 pm UC) to answer your questions, ask us anything!

---

How many other documents like this exist in museums, universities and private collections? Will you be able to spread the use of this technology to unlock more writings from the past?

Great work!!!!

microcosm315
William Brent Seales (WBS): There is no complete inventory world-wide of eligible objects, but from my own work I know there are lots of interesting objects. Almost every library/museum has damaged things. I know the IAA has some really intriguing pieces that we hope to work on next.

I would like to know from Seth Parker:

- how much of the software was written in-house?
- are there any plans to release it under an open-source license so that other research institutes can also use it?
- are there any links that explain how the software algorithms work.

And congratulations on a fine piece of work. I am extremely impressed.

UnlikelyExplanations

SP: We use OpenCV, ITK, and VTK for data structures and stuff, but all of the really important algorithms were written in-house by our students. We will release this open source next year as part of our NSF grant release. Right now, the methods section of the Science Advances paper (linked in OP) is the best place to learn about the algorithms used on En-Gedi, but we’ve got a lot more in the pipeline we’ve yet to describe.

As a Byzantinist, I want to first say thank you for this incredible work!

Do you have any plans to tackle some of the scrolls from Herculaneum and Pompeii? One of my professors once told me that there are a number of similarly "unreadable" scrolls from libraries at those sites, which if we could read them would almost surely contain some otherwise unknown ancient sources.

Porphyrius

Seth Parker (SP): Herculaneum was our first passion and what got us thinking about this idea of virtual unwrapping in the first place, so absolutely!

My question relates more to the technology that you developed to read the scrolls. Do you see it moving from a research lab to a field lab in the reasonable future due to the software naure, or is mainly tied to large pieces of equipment that cannot travel to the field.

Also, what other applications of the technology have been considered/explored? Do you believe the main benefit will stay in archeology, or do you think it will have a much wider effect?

Iron_Skin

SP: The largest part of this setup is the volumetric scan. We used a desktop-sized micro-CT scanner from Bruker Micro-CT, but this step can be replaced with a lot of stuff, from a desktop machine to a synchrotron. As scan methods improve, so will the ability to take things into the field.

The computational part currently can be run on any *nix system, great or small (though you probably want at least 16gbs of RAM and a Core i7 or equivalent). We really wanted to target normal computers that textual scholars would already have, so that’s what we did!
How you feel about the mixture of ancient history and, basically, sci fi? You've got to have had a couple of surreal moments like "guys, are we actually doing this?"

evilbrent

WBS: You are not kidding. When I look at that final image we produced, without rehearsing all the technical steps, it seems like a dream. Not sure how else to describe it - magic isn't the right word but the imaging is just so powerful. And I love the interdisciplinary part. As a computer scientist I have learned so much about history and other areas of scholarship.

What other historical artefacts might your team turn your attention to next?

Also, what concerns, if any, do you have moving forward to unravel the rest of these scrolls?

Sulauk

SP: Herculaneum was our first passion, so those scrolls for sure, but honestly we'd work on anything damaged. Every new document is potentially a new technical challenge.

The other En-Gedi fragments look exactly like used charcoal bricks, so I can't even imagine what their internal structure looks like...

What other historical artefacts might your team turn your attention to next?

Also, what concerns, if any, do you have moving forward to unravel the rest of these scrolls?

Sulauk

MS: I am hopeful that we will get images of further sections from the En-Gedi scroll itself.

There are also some other Dead Sea scrolls at the Israel Antiquities Authority which could possibly benefit from this technology.

I have a couple questions!

How long does a digital analysis like this take? How much of the work is manpower compared to the technology just doing its thing? Were there any unexpected challenges that came up during the project?

This is all really interesting thanks for taking the time to do an AMA.

Seekayem

SP: In this case, it took about 1 month to know we had something, another 5 months to know we had Leviticus, then another 6 months to complete the full scroll (Then another 8 months to write the paper and get published!). That was all while developing and improving the entire process. If I did the same work now, working 8 hours a day, I could probably redo it all in about a month.

How scary is it to work with something like that? What precautions are taken to avoid any damage?

Hairy_Viking

WBS: The curators do all the work in handling the objects, and they are experts. I've been amazed at
how careful and protective they are at moving things around and handling them. But I’m convinced that this is one of the safest kinds of analysis you can do, short of “doing nothing.”

how was this even recognized as a scroll? I looked up a picture of it. it just looks like a piece of charred stick

cubosh

WBS: It looks pretty badly damaged. But the archaeologists (Sefi Porath actually pulled it from the site - he’s a co-author on our paper) know with amazing accuracy what things look like and he told me they knew right away that the fragments could be manuscripts.

Have you had any luck getting access to the Vatican archive? If so will you able to make your findings public?

supabruda

SP: If we had and we can’t, would we be able to say so?

But seriously, if you know somebody with access, I’d scan anything with that much potential knowledge.

Is translated text available anywhere yet?

mrhone

Michael Segal (MS): The text of the two chapters of Leviticus that have been uncovered so far is letter-to-letter identical to the Masoretic text that we know from medieval manuscripts. Similarly, in these two chapters, the section/paragraph divisions are in identical locations to the medieval MT (choosing MS Leningradensis as our basis for comparison). Therefore, there will not be any differences in translation, since the text is identical to that which is generally translated in Jewish and Christian Bibles.

So the scrolls are read without unrolling them. That means you have to not only identify the ink but locate it in 3 dimensions then virtually flatten it out. Could a similar technique be applied to other ancient artifacts like the Antikythera mechanism?

unkunked

SP: We’re currently focused on page-like materials (e.g. books, manuscripts, scrolls, etc.) but the principles can be applied in a lot of places. It really comes down to exactly what features of the object you want to look at. There’s a really great article from The New Yorker about what works been done on the Antikythera mechanism.

Really interesting work!

How much manual work and contextual knowledge (if that's the right term) is required to digitize the scrolls? How far away are you from a fully automated process if someone wants to employ this on a large scale?
WBS: The scroll from En-Gedi was digitized/scanned with a commercially available scanner. Quite a lot can be done with available equipment. The software for analysis isn’t there yet, that’s where we’ve spent a lot of time.

Hey guys! I think the work you are doing is so cool - thank you so much for your contributions to the field!

A few questions - would be great if you could answer any or all!

In your work, what is the most interesting tidbit of info about the inhabitants you have come across?

I imagine you are or will be involved with many apocryphal texts that most people will not have seen. Is there anything that you have seen or read that has affected your personal beliefs?

Brent - What has been the biggest non-technical eye-opener for you?

Michael - What or how do you think these technological advancements will affect the modern world’s understanding of religion? Do you think that the internet / having open access to religious source documents leads to a reinforcement of religious institutionalism, or a de-centralization?

Seth - Do you see a lot of personal bias and inflection when working professionally on these things? Are there people from all ends of the religious spectrum (atheist, devout etc) who work on these projects, or does it tend to be a pretty homogenized group of believers? Do conversations about faith ever come up at work? Do they go pretty smoothly? (I imagine it may be a bit less of a hot-button topic for people in the field)

Thanks again fellas!

ChildofAbraham

SP: For my part, I've seen no bias. It's a pretty diverse field, so everyone has their own interests at heart, but what we share is the desire to just learn more. And honestly, I want to enable everyone to learn as much as they can, personal philosophy and religion aside.

Hey guys! I think the work you are doing is so cool - thank you so much for your contributions to the field!

A few questions - would be great if you could answer any or all!

In your work, what is the most interesting tidbit of info about the inhabitants you have come across?

I imagine you are or will be involved with many apocryphal texts that most people will not have seen. Is there anything that you have seen or read that has affected your personal beliefs?

Brent - What has been the biggest non-technical eye-opener for you?

Michael - What or how do you think these technological advancements will affect the modern world's understanding of religion? Do you think that the internet / having open access to religious source documents leads to a reinforcement of religious institutionalism, or a de-centralization?

Seth - Do you see a lot of personal bias and inflection when working professionally on these things? Are there people from all ends of the religious spectrum (atheist, devout etc) who work on these projects, or does it tend to be a pretty homogenized group of believers? Do conversations about faith ever come up at work? Do they go pretty smoothly? (I imagine it may be a bit less of a hot-button topic for people in the field)
Thanks again fellas!

ChildofAbraham

WBS: I’ve been surprised at the conservation community being less enthusiastic about trying these methods. The risk is quite low, and the payoff can be huge in terms of discovery. I've also been surprised at the difficulty I've had to get funding, but I suppose that's an ongoing concern for everyone doing research in these economic times.

Hey guys! I think the work you are doing is so cool - thank you so much for your contributions to the field!

A few questions - would be great if you could answer any or all!

In your work, what is the most interesting tidbit of info about the inhabitants you have come across?

I imagine you are or will be involved with many apocryphal texts that most people will not have seen. Is there anything that you have seen or read that has affected your personal beliefs?

Brent - What has been the biggest non-technical eye-opener for you?

Michael - What or how do you think these technological advancements will affect the modern world's understanding of religion? Do you think that the internet / having open access to religious source documents leads to a reinforcement of religious institutionalism, or a de-centralization?

Seth - Do you see a lot of personal bias and inflection when working professionally on these things? Are there people from all ends of the religious spectrum (atheist, devout etc) who work on these projects, or does it tend to be a pretty homogenized group of believers? Do conversations about faith ever come up at work? Do they go pretty smoothly? (I imagine it may be a bit less of a hot-button topic for people in the field)

Thanks again fellas!

ChildofAbraham

MS: I think that the more information that is open access and available to all, the more people have the option to choose what to believe or not. And everyone should have access to make informed decisions.

What is one piece that you would love to get your hands on?

jbhilt

SP: There's a damaged edition of Seawolf (Jack London) at the Huntington Library that would be really interesting to work on. Also, the Franklin Expedition papers.

Are there any surprises in the text? If not in the content, then in the use of language or the handwriting?

ohdon

MS: The "surprising" feature of the En-Gedi Leviticus scroll is how similar it is to the medieval Masoretic text, matching its consonants and paragraph divisions in all of its details. We know of other early scrolls that are nearly identical, but this is closest that we have seen to date (at least in the
columns that have been developed until now).

Have you ever played a prank where you photoshop your own rendering to show to people and draw like, spaceships on it or references to Atlantis or something?

**XSplain**

WBS: the idea of pranking certainly has a temporary appeal, but the overarching idea of "getting it right" and making sure that there is a provenance chain from the scan all the way to the final "texture claim" is very very important. As Emanuel Tov reminded me, "this is important stuff" and I agree with him 100%.

Have you ever played a prank where you photoshop your own rendering to show to people and draw like, spaceships on it or references to Atlantis or something?

**XSplain**

SP: I'm a pretty big fan of shopping things for a joke, but never with research stuff like this. People seem to think you're crying wolf when it comes time to publish a paper like this? Can't imagine why...

Who funds your research? Who makes money off of this sort of thing?

**Jowitness**

WBS: The National Science Foundation is funding this work, as well as Google. We’re not doing the work as a money-making enterprise, but as scholarship. There are so few manuscripts from antiquity available - anything adding to that corpus is really valuable and interesting. There are broader impacts from the work - the algorithms we’re building have application in other areas too.

Sounds amazing, congratulations!

Can you also see illustrations? And do you plan on publishing scans like this in a coffeetable book (this is a request).

**Decestor**

SP: Our software is highly tuned to page-like things, so anything on a "page" can theoretically be seen. So if you made an illustration with a lead-based ink, absolutely! If you used a carbon-based ink, maybe, maybe not. That remains to be seen...

The obvious question is: can you use this technology to read scrolls from the library at Herculaneum?
But also: could it be used to read Mayan books found in graves that are now solid lumps of limestone?

tchad49

SP: This is one step towards both of those goals! Now that the framework is in place, the key is tuning each step in the framework to the specific challenges that each of those texts hold. I'd love to know more about these Mayan books, btw. Is there a good write-up somewhere?

The obvious question is: can you use this technology to read scrolls from the library at Herculaneum?

But also: could it be used to read Mayan books found in graves that are now solid lumps of limestone?

tchad49

WBS: I have always believed we will read Herculaneum non-invasively. We're closer now than ever.

What did it say? Can I read it somewhere?

couldntgive1fuck

MS: The first two columns of the En-Gedi scroll are identical to the medieval Masoretic text (with respect to consonants and section divisions). For a transcription of the text, see our article in Textus (the journal of the Hebrew University Bible Project), found at the following link: http://www.hum.huji.ac.il/upload/_FILE_1474480873.pdf

What's the runtime for these computational tasks? What kind of hardware?

EACCES

WBS: The data sizes are huge. Off the scanner we're talking about 4kx4k slices, and maybe 10k of those. So the early processing of the volumes is time consuming. We run on multithreaded workstations with lots of memory. Once we get going on the segmentation it's easier because we have localization in the volume. For the ML stuff we're doing now (we didn't use it for En-Gedi), we're using NVIDIA graphics cards.

As someone with an interest in software development and data analysis for multiple different types of microscopy/spectroscopy/tomography, Can you mention a bit about the computational side of things for this project?

Did the analysis use pre-existing software or was it specifically written for this one exact project?

What languages were used for the software/analysis routines?

Are there plans to make this type of software open source or is this proprietary? Thanks again!

justphysics

SP: We used a number of libraries (ITK, VTK, OpenCV) to facilitate stuff we were doing, but all the heavy-lifting was done by custom software, written in C++. It will become open source at the end of our current funding cycle (sometime next year).
Can you use this technology to, say, "unfold" and reveal tattoos in ancient mummies?

sororytosaythis_but

SP: I've never thought about that before, but yeah, why not? Ink written on parchment (animal skin) is actually kind of a similar thing when you think about it, and that's exactly what the En-Gedi scroll is!

How long did it take you to complete this process? Have you now learned enough that it could be more readily applied to further artifacts, and how might that rate of deconstruction improve? What "performance bottlenecks" are there in the process (e.g., would faster computers, or more detailed scanning help)?

gwinerreniwq

SP: It took about a year from the dataset making it to our lab to a complete unwrapping. What we have now is a pretty solid pipeline, each stage of which can be independently improved. In the past year, we've more than doubled our productivity, but there's still a lot more to be done. And it's a really exciting time in computing, so there's currently no limit on what could happen!

A note about scanning resolution: the higher the resolution, the larger the dataset, the longer the processing times. Calibrating the optimum resolution for any given object is object dependent and will probably always be a little bit experimental.

Will this technique be any use on suspected palimpsests?

zed1207

WBS: Other excellent techniques exist for palimpsests, including spectral imaging and x-ray fluorescence. Amazing work by Mike Phelps at St. Catherines and William Noel on the Archimedes palimpsest.

Will this technique be any use on suspected palimpsests?

zed1207

SP: Maybe! The hard part with palimpsests is, of course, there are multiple layers of ink. Having a 3D representation of the document might help, but it really depends on having the right type of 3D representation.

How close is the King James version of Leviticus to the En-Gedi?

aVarsityLetterman

Michael Segal (MS): The text of the two chapters of Leviticus that have been uncovered so far is letter-to-letter identical to the Masoretic text that we know from medieval manuscripts. Therefore, there will not be any differences in translation, since the text is identical to that which is generally translated in Jewish and Christian Bibles.
How close is the King James version of Leviticus to the En-Gedi?

aVarsityLetterman

MS: The text of the two chapters of Leviticus that have been uncovered so far is letter-to-letter identical to the Masoretic text that we know from medieval manuscripts. Similarly, in these two chapters, the section/paragraph divisions are in identical locations to the medieval MT (choosing MS Leningradensis as our basis for comparison). Therefore, there will not be any differences in translation, since the text is identical to that which is generally translated in Jewish and Christian Bibles.

What inspired you to begin this work?

luckyphoenix34

WBS: The digital library movement got me interested, back in the 90’s, but I quickly got into the damaged part of materials. Digitization is crucial to build a library. What do you do if something is really hard to digitize? That was the initial thinking. Then when I thought about the most profoundly damaged things it took me right to the idea of “imaging without touching / opening.”

What other ancients texts are left to find?

iamtherealmrb

SP: The thing is, we're trying to read an invisible library. No one really knows how much stuff is out there because no one can currently read it to catalogue it!

What other ancients texts are left to find?

iamtherealmrb

MS: When the Dead Sea scrolls were discovered in the 1940’s, they included many texts that we already knew, but also compositions that were previously unknown. There is no way to predict this in advance. We will only know when we find them.

I know that this isn't exactly your field, but can you summarize the current theories on the Scroll origins? I always grew up hearing that they were written and stored by the Essenes, but recently I've read that many scholars reject that hypothesis and there's a theory that the scrolls were from many different generations of groups that lived in Qumran.

Are any of you religious daati? How do you view your studies through a religious lens, if at all?

As a side note, I should be taking more interest in the Scrolls. My father's cousin was Mordechai Breuer, who was also a Biblical professor at Hebrew U.

Yserbius

MS: As you noted, the question of the origins of the DSS are beyond the discussion here. However, just to state briefly - the Qumran caves (the main site in the Dead Sea area) contain over 900 scrolls, about 200 of them biblical manuscripts. Among the over 700 remaining manuscripts, some can be classified as "sectarian" scrolls, while others are almost certainly of other origins. Think of it more as a collection of scrolls from the whole Land of Israel (and maybe outside too), which different people...
brought with them when coming to Qumran. That explains the diversity of the texts there.

I am a religious Jew - the question of the relationship between religion and critical inquiry is a fascinating topic, but far beyond the discussion here.

Please continue your interest in the Dead Sea scrolls!

What religion are you and did this project increase or decrease your faith in God?

**IBuyHousesTX**

MS: I am Jewish. This discovery is a great example of the possible symbiotic relationship between the study of science and religion.

What religion are you and did this project increase or decrease your faith in God?

**IBuyHousesTX**

WBS: I am a Christian and the discovery / recovery of the text within the En-Gedi scroll has helped me upgrade my understanding of these ancient texts and their place in the development of both Christianity and Judaism. In my case faith and devotion is rooted in reason and “all truth is God's truth.”

So what’s your next delicate, damaged materials project?

**mollywoppie**

SP: There’s a lot of things in the works, but we are still and always will be dedicated to reading a full Herculaneum scroll.

I’ve found articles about that from 2015, how come you’ve analyzed it just yesterday? (or has 2015 been just the beginning of the analysis?)

**Yegor40**

WBS: In 2015 we reported the identification of the text (Leviticus) but we hadn’t completed the technical work or the software pipeline. It took us until now to build the software, apply it properly to the data, write the paper, get it accepted, and prepare all the data for release. By the way all the original data is available (link is in the paper).

How many pages deep can your tech go? and thanks.

**noeljb**

WBS: The materials with writing aren’t that dense (papyrus, leather). So tomography can go through many layers. The Herculaneum scrolls have 100's of wraps, for example. Xray based methods are amazing.

As a North American archaeologist, I will never not be jealous of the absolute trove of artefacts and features you old world archaeologists have.
WBS: Right? That treasure trove was compelling for me also, even as a computer scientist.

Are we at the point where such things should never be attempted to be analyzed in the old-fashioned methods?

Robert_Cannelin

SP: Probably not yet, but it's certainly a step in that direction. We hope that this starts getting curators and conservators thinking about alternative preservation methods they maybe haven't considered previously!

I understand what was done by the team. That was very hard to do. Bravo!

Does the team foresee similar mapping/rendering capability for noncarbonised scrolls, such as, could this technique be used to similarly scan and "unravel" the mysteries on brittle papyrus scrolls with faded ink?

What about curled up metal scrolls like the small ones found among the DSS?

Does the team foresee the ability to help render an unrollable palimpsest scroll so both texts can be studied?

When many things are crammed together into a genizah, the various parchments and other media tend to get stuck together. Can this technology sort their literary lumps, like the Cairo genizah bits that still haven't been studied?

There are many ancient scrolls and codices that are in various stages of decay and, frankly speaking, they all need to be correctly rechecked and scanned to have a record of them before the ravages of time get to them as well.

Does the team foresee an ability to "3D scan" and save all most ancient inoperable objects and manuscripts through software and equipment portability for the legacy of the human race?

Lulwafahd

SP: The En-Gedi scroll was burned and not carbonized, so the answer to your first question is yes, absolutely! The reason we did a pipeline and not one algorithm for virtual unwrapping is because you need to be able to configure each step to optimize for the object you're working with. For example, ink-based texts show text as visual (luminance/color) contrast, but metal scrolls with etchings show text as morphological contrast (a height difference within the local area). So you tune the texturing portion of our pipeline to figure out that morphological contrast, and the rest stays the same.

So real life dead sea scrolls? Prep for impact boys.

BrokenAscendent

WBS: It is a researcher's dream - impact.

How does it feel to be perched at such an interesting point of history, with the technology to read such a scroll digitally, when it would have previously been impossible?
Do any of you have regrets about similar documents that were inevitably destroyed when some discoverer attempted to open and read the contents?

MystJake

SP: Exhilarating! Sometimes it feels like only other people are making real changes in the world, then something like this happens and you realize that everyone is having some sort of impact, great or small. I feel blessed to have been involved in something so huge!

I always regret the loss of information and I hope people realize that we haven't moved past this sort of destruction. Very important manuscripts, religious or otherwise, are being destroyed today. Sometimes it's unintentional through conservation-gone-wrong, but there are many, many people who are doing so intentionally.

Dit you do the computing on Linux? If yes, on what distro? And did you have a specific reason for choosing Linux or that specific distro?

Jarco5000

SP: Everything's written in C++ and built using cmake. We actively develop on OSX, Ubuntu, and Debian, but there's no reason it shouldn't run on other distros. It currently probably wouldn't build on Windows, but that's mostly because of system specific stuff (file paths, system libs, etc).

What kind of materials did you practice on to give you confidence your machine could accurately digitize hidden things?

BrujaBean

SP: We've been thinking about and working on specifically virtual unwrapping for about 15 years. During that time we've done many, many lab-created examples to make sure we were doing things correctly.

What was the main focus of the scrolls content?

Scoutabout315

MS: It contains material from Leviticus, chapters 1-2.

This is the most exciting tech news I've heard in a month! If you can scan a scroll, can you scan a book? Would the technological challenges be similar?

sempf

WBS: Books have their own challenges. We've scanned some and yes we can see writing. But the working volume of the scanner has to be larger for a book. Also tomography is not as great when an object is more flat than it is spherical - x-rays have to pass through from all orientations. There is a frontier with re-used manuscripts that are inside of book bindings, and this and other methods for imaging them are emerging.

How does what you've uncovered differ (if it differs) from our currently translated copies of Leviticus?
Michael Segal (MS): The text of the two chapters of Leviticus that have been uncovered so far is letter-to-letter identical to the Masoretic text that we know from medieval manuscripts. Since most of our translated copies of Leviticus are based upon the Masoretic text, then it is essentially identical to this scroll.