

# The Future of the Past

Textual scientist Gregory Heyworth lights up the first drafts of history.

By Kathleen McGarvey

Pack lightly, seasoned travelers advise. Take only what you need.

And Gregory Heyworth, an associate professor of English, does. A scant collection of clothes makes it into his bags when he flies to Italy, or the former Soviet republic of Georgia, or Wales. He pares his wardrobe to make room for the camera, panels of LEDs, computers, and other pieces of equipment that fill his luggage instead.

Trained as a scholar of medieval literature, Heyworth has become—in a term he coined—a “textual scientist.” He recovers the words and images of cultural heritage objects that have been lost, through damage and erasure, to time.

He calls it “forensic science applied to literature and to history.” And while he paints the effort in swashbuckling terms—he has described himself as “an adventurer in an undiscovered country, searching for the hidden text”—a somber urgency drives him. At a bare minimum, he estimates, there are 60,000 manuscripts from before 1500 in Europe alone that are damaged to the point of illegibility.

Some of the objects are faded; some, charred; some, moldy and crumbling. Some surely hold secrets that could change our sense of history. And time is not their friend.

To rescue them, Heyworth and his collaborators on the aptly named Lazarus Project created the contents of his airline baggage: a transportable multispectral imaging lab—the only one in the world—that uses different wavelengths of light to photograph cultural artifacts. The team analyzes the images, digitally salvaging ancient manuscripts, maps, and other texts too delicate and precious to transport. They make the undecipherable, and even the invisible, legible again.

To do so requires a broad amalgamation of expertise—and much of that breadth is represented by Heyworth himself.

Michael Phelps, director of the California-based Early Manuscripts Electronic Library (EMEL) and a frequent collaborator, calls him an “engine of innovation.” Textual science, he says, couldn’t exist





Handwritten text in a medieval script, likely Georgian, with a large, ornate initial letter at the top of the page. The text is arranged in several lines, with some words appearing to be in a different script or dialect. The page is decorated with a large, colorful initial letter and a decorative border at the top.

ILLUMINATING: Textual scientist and English professor Gregory Heyworth uses multispectral imaging to digitally restore ancient manuscripts, like this 12th-century manuscript of the Gospels, now held in the Svaneti Ethnographic Museum in Mestia, Georgia.



without someone like Heyworth. “It needs someone who is a humanities person, who knows the manuscripts and the collections. And someone who has some chops in hard science, too—someone who brings both sides of the equation together.”

The work requires expertise in paleography—the study of old handwriting—and codicology, the study of books and manuscripts as material objects and cultural artifacts. It demands proficiency in imaging science and computer science. It calls for knowledge of chemistry, to understand the makeup of inks and their interactions with the surfaces on which they were used.

The cadre of people able to carry out the projects is vanishingly small—which is why Heyworth sees one of his most important tasks as developing a textual science curriculum to train students.

“There are only about 20 people in the world now who do this,” he says. “And that has to change.”

Heyworth joined the University in July, from the University of Mississippi. Henry Kautz, the Robin and Tim Wentworth Director of the Goergen Institute for Data Science and a professor of computer science, spotted him when he led a search committee seeking to hire faculty whose nontraditional research could bridge different disciplines. “It’s state-of-the-art work,” says Kautz.

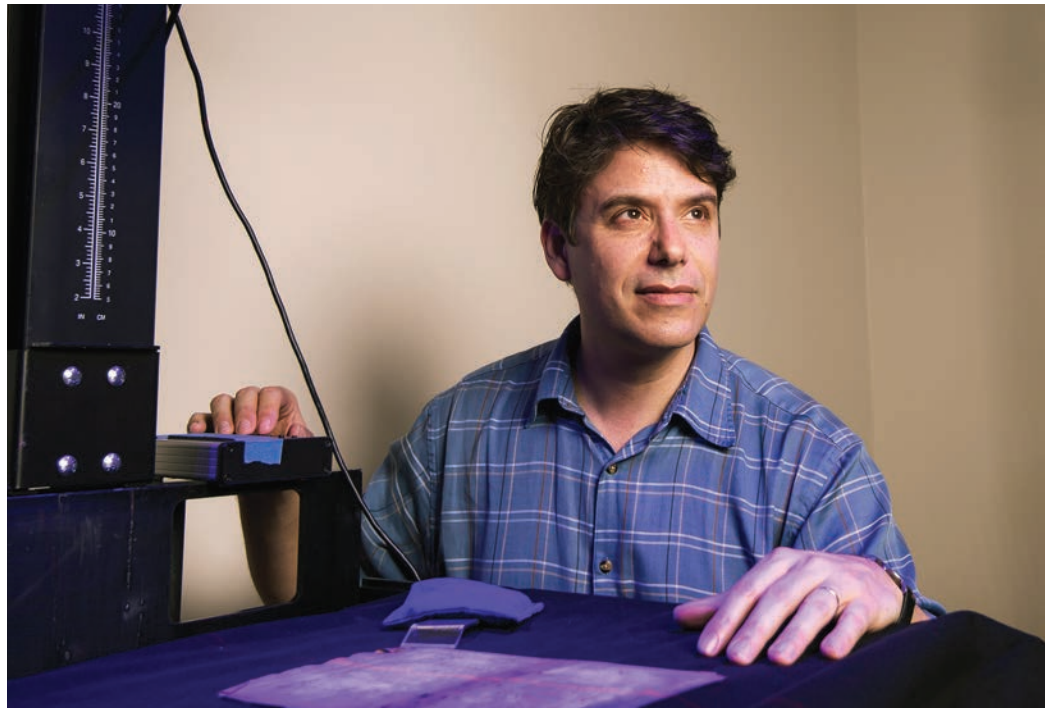
Heyworth arrived at his specialty “by necessity,” he says. After

earning a doctorate in comparative literature at Princeton, he spent five summers in Dresden, Germany, scrutinizing “Les Eschéz d’Amours,” (“The Chess of Love”), an anonymous, 14th-century poem that’s one of the longest verse works in French. Only a handful of manuscript copies ever existed, and just two survived: a fragmentary copy in Venice and a nearly complete copy, lacking only the ending, in Dresden. But that manuscript was virtually destroyed by water during Allied bombing in 1945, “leaving the manuscript’s once pristine parchment a faded and murky Rorschach of figures on blotting paper,” Heyworth has written.

Though scholars considered the poem to be effectively lost, Heyworth refused to give up. For those five years, he used ultraviolet lamps to try to recover traces of the writing, a long-established salvage technique first developed by a German monk in 1914. He followed the method as far as it would take him—but it wasn’t far enough.

Determined to find a way forward, he began to look online for ideas. And that’s how he discovered the work of Rochester Institute of Technology (RIT) imaging science professor Roger Easton, who had led a team in using multispectral imaging to recover works by Greek mathematician Archimedes from a 12th-century palimpsest.

Palimpsests are essentially recycled pieces of treated animal hide, better known as parchment. Parchment was a precious commodity,



**TEAMWORK:** English graduate students Helen Davies and Alex Zawacki (left) examine a 15th-century manuscript binding fragment under the multispectral imaging equipment in the lab of Gregory Heyworth (above), an associate professor of English. Training new experts in his techniques is a vital task for Heyworth.

and so ancient and medieval scribes would reuse it, literally scraping away old writing to create a blank surface for new text.

In the case of the Archimedes palimpsest, seven of his treatises—10th-century transcriptions of his work from the 3rd century BCE—had been erased and the pages rebound and written over as a prayer book by monks. An anonymous buyer purchased the piece—called a codex, an ancient manuscript assembled in book

form—at auction at Christie’s New York in 1998. He entrusted it to what is now the Walters Art Museum in Baltimore for conservation, an effort that spanned a decade. Imaging and analysis by Easton, physicist and imaging specialist William Christens-Barry, and Keith Knox ’70, ’75 (PhD), formerly an imaging scientist at Xerox, was fundamental to an effort that yielded about 80 percent of the original text and revealed a previously unknown work by Aristotle.

“The flash of a camera and a clever algorithm . . . can transform a page,” wrote Walters curator William Noel of the process, in a book he coauthored with Reviel Netz, *The Archimedes Codex: How a Medieval Prayer Book Is Revealing the True Genius of Antiquity’s Greatest Scientist* (Da Capo Press, 2007).

Here, Heyworth realized, was the key that might open the “Eschéz” manuscript.

He contacted Easton in the summer of 2009 with a plan. Heyworth applied for—and received—a grant from the National Center for Preservation Technology and Training to build a transportable version of the multispectral imaging lab that Easton’s team had used in Baltimore. Easton, imaging entrepreneur Ken Boydston—whose company, Megavision, created the camera Easton used—and Michael Phelps worked with Heyworth to create the system.

What the naked eye can detect is only a fraction of what

manuscripts—or maps, or globes, or inscribed stones—contain. But multispectral imaging can expose what otherwise lies hidden.

The method developed by the Archimedes group relied on remote sensing technology invented by military and environmental scientists to photograph terrain using 12 light frequencies between ultraviolet and infrared light. The innovation of Easton’s team was using light-emitting diodes, or LEDs, as the light source. Cool LEDs don’t subject manuscripts to the damaging heat that broadband light does, and they bring greater efficiency, exposing pages only to the desired form of energy. The method “produces much better results and does no damage to the object,” says Heyworth.

He met Easton for the first time in 2010, when Easton and Boydston delivered the system to him at the airport in Dresden. Their efforts in deciphering the damaged “Eschéz” manuscript were successful. About 96 percent of the original text was recovered; Heyworth and coeditor Daniel O’Sullivan have published the first volume of a two-volume critical edition of the poem, with the second soon forthcoming. “Les Eschéz d’Amours” is now the last known major medieval allegory to have found its way into print for modern readers.

**T**he project was the start of an enduring collaboration between Heyworth and Easton, both through the Lazarus Project and now through the freshly formed group RCHIVE, which takes advantage of their new physical proximity and draws together faculty and students at the University and at RIT. Under the auspices of the Lazarus Project, Heyworth, Easton, and a team of colleagues and students have already worked with such important manuscripts as the Vercelli Book, the oldest book of English; the Black Book of Carmarthen, the oldest book in Welsh; and some of the earliest Gospels, now held in Tbilisi, Georgia.

“Greg has much better blue-sky vision than I do,” says Easton of Heyworth, “to the point where I have to say, ‘no, we have to focus on what we do well, and get somebody else to do that. We have enough of our own work to do.’ But it doesn’t affect him. He wants to do it anyway.”



**ON SITE:** Heyworth (above, second from left) and colleagues Michael Phelps (left) and Roger Easton examine materials at the Archive and Capitulary Library of Vercelli, Italy, as head curator Timothy Leonardi (right) looks on during a 2014 visit. The team often sets up in spaces that date to the Renaissance and Middle Ages (opposite).

A covered bridge in Bucks County, Pennsylvania, beckons—it holds graffiti that may be related to a visit to the area from Abraham Lincoln. The Dresden collection that holds the “Eschéz” manuscript is also a vast treasure house of baroque music—much of which has been in some way damaged. Heyworth did a pilot project there last summer, imaging music by composers such as Antonio Vivaldi, Georg Philipp Telemann, and Carl Ditters von Dittersdorf. “There are three unknown operas there, damaged, that we are recovering, and then symphonies by important composers of the time, like Johann David Heinichen.” Now that he’s at Rochester, Heyworth hopes to collaborate with musicians at the Eastman School of Music to take the manuscripts from recovery to performance.

“Sometimes when letters are erased, they leave a ‘footprint’ behind,” he says. Even when the ink has only flaked away, there’s still a trace. The acid in the ink eats away at the parchment, creating a channel where the parchment is thinner. Lighting the manuscript from beneath allows the team’s camera to capture what was lost—the thinner channel transmits light better than the parchment that wasn’t written on, and lost words can shine out more brightly than their surroundings.

“We’re seeing back in time. That’s really what we do,” says Heyworth. “We’re seeing back before the final draft. I like to say that our job is to try to recover the first draft of history, before people started changing things. And you know how revealing that can be—all the changes of mind, the attempts that are erased and started over again. That gives us huge insight into what the author or cartographer knew, understood, and aspired to.” In that sense, the process can recover not only what was once on the page, but also the story behind its creation.

The transportability of the lab is crucial. Multispectral imaging systems are scattered around the world, at places like the Library of Congress, the British Library, and the Israel Antiquities Authority. But they don’t travel.

And most of the manuscripts in jeopardy are held by institutions that lack financial resources and that can’t, for various reasons, send their objects to establishments with digital conservation equipment.

“Transportability was fundamental,” says Timothy Leonardi, the head curator of the Archive and Capitulary Library of Vercelli in Italy. “It’s not possible for us to bring our manuscripts outside the library.” When Heyworth and his team visited, twice, to image and analyze items from the collection, they carried out all of their work in the library, using the base of an 11th-century tower, and a suite of rooms formerly used by Pope John Paul II, as their center of operations.

Easton describes the sight of Heyworth toting the lab to the project: “I left Greg at a train station in Rome, and he had this little train of bags, tied together with a webbing, that he was pulling. He had the camera and the copy stand and the computers and the lights. And so it’s still not portable in a convenient sense. But you have to take it places. The manuscripts don’t leave where they are, and we wouldn’t want them to leave. We don’t even want to touch them.”

**T**he Lazarus Project is a not-for-profit organization that seeks to provide the team’s services at little or no cost to individual scholars and smaller institutions around the globe. It claims no ownership of the images and analysis produced, and diplomacy is central to its process.

Heyworth formed the Lazarus Project as a separate charitable group that only affiliates itself with the university at which he’s employed. That separation is important, he says, in addressing concerns about a university’s appropriation of the collections he works with. “With the Lazarus Project, we’re able to tell them, we don’t own your data. Not only do we not own it, we never want to own it. It needs to be published freely, openly—and that’s what we’ll do. In return for doing it for free, we will open up this to everyone.”





The work is painstaking. It takes two days to set up and calibrate the system. Michael Phelps, Heyworth's collaborator, ticks through some of the preparatory questions: "Are you dealing with flat pieces of paper or parchment? Can you just set them down on a copy stand underneath the camera? Are you dealing with a codex? If you're dealing with a codex, is it intact? Is it broken? How fragile and brittle is it? How much can it be opened? Can it be opened 100 degrees? Or 90 degrees? Or 150 degrees? You have to carefully plan out how you're going to handle the object."

The room is readied, the windows darkened with black aluminum foil or black rubberized cloth. Banks of lights are put in place to illuminate the pages from above and below.

And then the imaging begins, with a camera whose lens is made of quartz. There are fewer than a dozen such cameras in existence.

When Heyworth and the team began their work in Dresden, they took 12 images of each manuscript page. Now, depending on how fragile the object is—and consequently, how efficiently it can be manipulated—they can image between 30 and 40 pages a day, with 30 to 50 images of each page.

Analysis of the images usually begins on site, so that archivists and scholars can give immediate feedback. And despite the elaborate technology involved, some parts of the process rely on startlingly

simple equipment—like a pair of comfortable shoes.

"We capture data onto a very fast external hard drive," Phelps says. "And then we 'sneaker-net' it, walking down the hallway to another room where image processing is set up." Networking the computers that capture images to those that process them, by recombining photographs taken at different wavelengths, slows down the process too much. "So we capture images for a certain number of hours, and then we walk them down the hall."

Heyworth was recruited to Rochester as part of the University's data science initiative. "In our projects, we generate an enormous amount of data, terabytes of data," he says. Data science—extracting meaningful information from large-scale data—is one of the University's top academic priorities for the next several years. The Goergen Institute for Data Science, headed by Henry Kautz, was formed in part to foster work that spans the entire University, linking—in combinations that vary by project—inquiry in science, medicine, the arts and humanities, social science, engineering, and business. To support such research and teaching, the University plans to hire new faculty in areas where data science plays a critical role.

The data on which Heyworth relies—images, optics, chemical analysis, and the algorithms to extrapolate information from the data—allow him to move beyond the parameters of traditional literary

# Illuminating the Past

Textual scientist Gregory Heyworth and his team use multispectral imaging to recover and preserve damaged and illegible cultural heritage objects, recovering text that has been lost. They photograph the object under a controlled spectrum of light. The images, when processed, allow the researchers to see material undetectable by the naked eye.

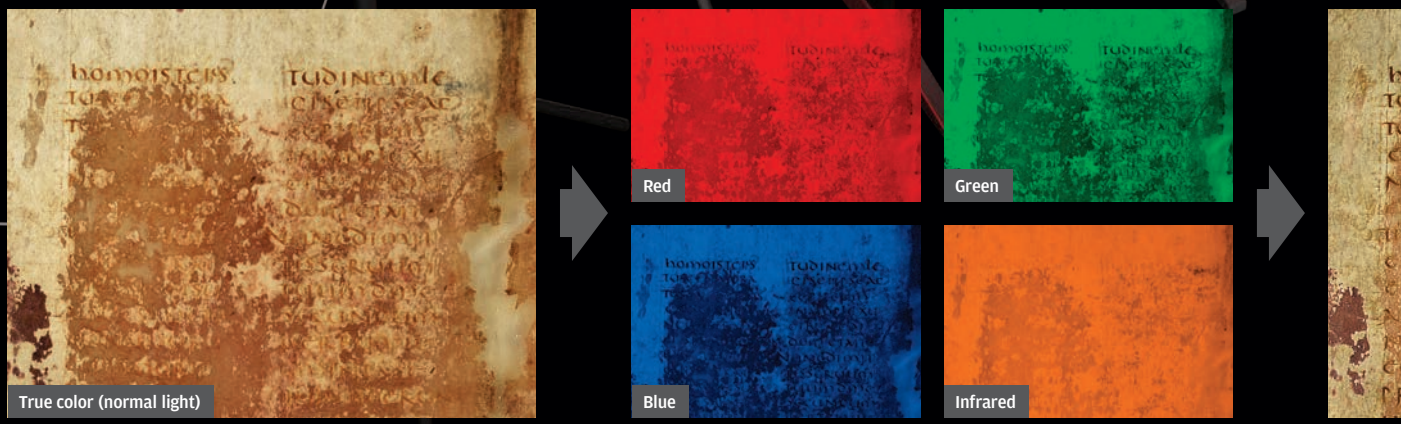
1000 nm  
900 nm  
800 nm  
700 nm  
600 nm  
500 nm  
400 nm  
300 nm

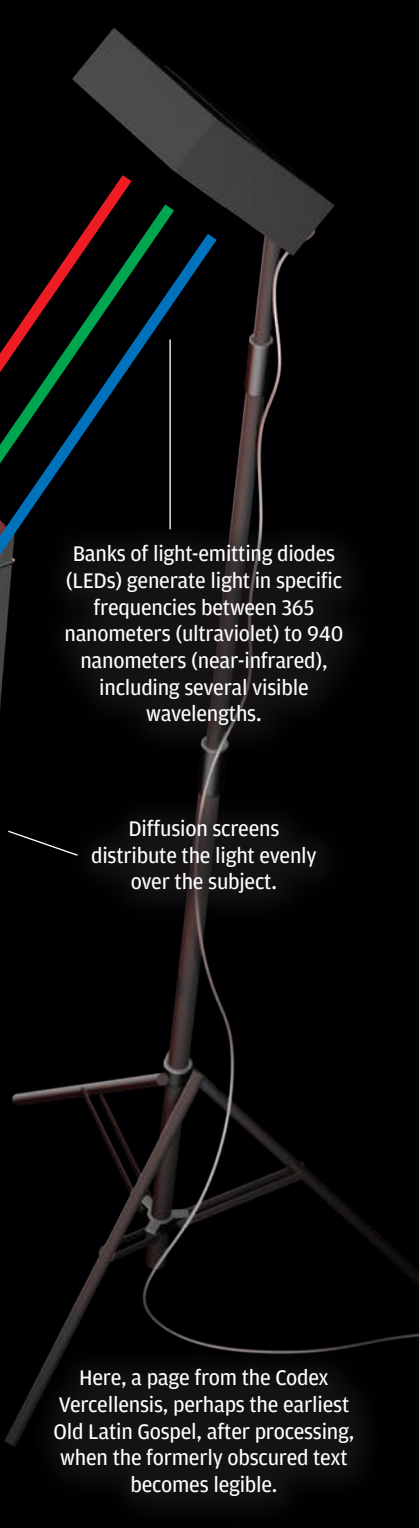
INFRARED  
ELECTROMAGNETIC SPECTRUM  
VISIBLE LIGHT  
ULTRAVIOLET

◀ 940 nm  
◀ 365 nm

A 50-megapixel digital camera captures a series of monochrome images in each of the selected wavelengths.

Dozens of images of each page are captured in various wavelengths. The LED lights may illuminate the manuscript from above, as shown here. They may also be placed at low, raking angles—to expose grooves left behind long ago by the scribe's pen—or below the manuscript so the light is transmitted through it. The images are then digitally combined and processed to enhance contrast and to show things that aren't visible in normal light.

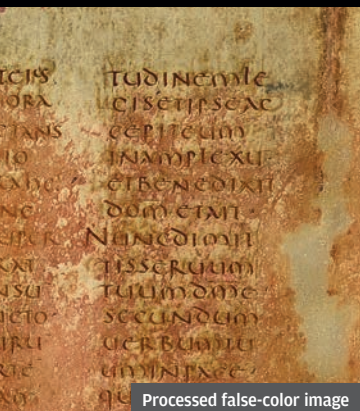




Banks of light-emitting diodes (LEDs) generate light in specific frequencies between 365 nanometers (ultraviolet) to 940 nanometers (near-infrared), including several visible wavelengths.

Diffusion screens distribute the light evenly over the subject.

Here, a page from the Codex Vercellensis, perhaps the earliest Old Latin Gospel, after processing, when the formerly obscured text becomes legible.



Processed false-color image

scholarship to a version of archival work that wouldn't be possible otherwise. At one level, that means recovering texts, like the "Eschéz," that have seemed irretrievable. At another, it means engaging with the shadowy history of a document that is revealed when it's subjected not only to the scholar's probing eye but also to the equally potent gaze of cameras and computers.

Says Phelps, of the Early Manuscripts Electronic Library: "Manuscripts are extremely complex objects. They have complex histories, where they've been damaged by fire or water, they've been degraded by mold and the mandibles of insects, and sometimes they've been intentionally effaced, either by painting over some of the text or by erasing text, in the case of a palimpsest. The chemistry of the inks is complicated and differs from manuscript to manuscript."

In the face of so many variables, the team's approach has been to "throw the kitchen sink at each manuscript," he says, trying different modes of imaging—including volumetric imaging, thermography, and X-ray fluorescence—and different wavelengths of light. And imaging contributes to material analysis, too, allowing the researchers to identify inks and pigments.

"We don't hunt and peck," Phelps says, searching through images to find the information worth processing. Instead they process massive quantities of data using statistics and mathematical algorithms. "Going from 12 images to 50 per page is a big jump, and that adds time to a project and it adds data that you have to store and manage later," says Phelps.

Now one of Heyworth's priorities is understanding why a particular technique works in a given situation, allowing the team to image and analyze objects more efficiently and effectively.

"This is where Greg's move to Rochester is important," Phelps says. "With the University of Rochester and RIT within a few miles of each other, we have the opportunity for hard scientists and humanities people at both institutions to put their heads together and create solutions. Heretofore, most of the solutions have been fairly ad hoc."

**B**e that as it may, the solutions have worked well enough to elicit at times what Lazarus Project board member and cartographic historian Chet Van Duzer has termed a "spectral gasp."

Van Duzer was part of the team from the Lazarus Project and EMEL who in 2014 imaged and analyzed the Martellus Map, a large—four by six-and-a-half feet—hand-painted paper map of Eurasia and Africa by German cartographer Henricus Martellus, who created it in Italy around 1491. Scholars believe that Christopher Columbus may have consulted the map before his voyage to the New World.

Acquired by Yale's Beinecke Rare Book and Manuscript Library in 1962, the map had hung, largely unnoticed, on a wall outside the library's reading room. Its inks had degraded with age, darkening the map until it became virtually undecipherable.

"It looks like a desert," Heyworth says. "It



**COLORFUL VIEWS:** The equipment—set up here in Vercelli—exposes manuscripts to different wavelengths of light, including blue, red, and ultraviolet, in an effort to recover as much lost text as possible.

looks dull green for the ocean and Sahara Desert—brown for the rest, and there's almost nothing that's visible."

When librarians gathered to see the team's before-and-after images, they let out a gasp, astonished to see for the first time what had been cloaked. There were places names all down the coast of Africa. For the first time, Japan was represented with a north-south orientation. The margins were decorated with descriptions of lands and people. The map's darkly impenetrable legend could suddenly be read.

"Details all over this map are now fully legible," says Phelps. "It's mind-blowing."

His is a mind not easily blown. "I've gotten sort of used to this, so I've gotten cynical. I've seen [the process at work] a thousand times, and you're not going to be able to get me to gasp very often." But working in Vercelli with the Codex Vercellensis—thought to be the earliest manuscript of the Old Latin Gospels—he felt a chill. "There are some pages where, if you're holding the manuscript in your





hands, you just don't think there's anything there. It's too far deteriorated. There's no technology other than a time machine that's going to recover it. And then—pow! We see legible text.”

But such recovery can only be carried out on anything approaching the needed scale if there are enough trained specialists. And that's why Heyworth is driven to enlist students, undergraduates as well as graduate students, in the effort.

He began teaching at Rochester this spring, with two courses—Image, Text, and Technology, and Digital Imaging: Transforming Real into Virtual. The second draws undergraduates and graduate students, but the first is an introductory textual science course aimed, as it was at Mississippi, at second-semester sophomores, so that they're ready by their senior year to carry out a major piece of work.

“They'll have the technical skills to be able to pull it off,” he says. “And they can present on it, or even publish on it, either individually or collectively.” For undergraduates, he says, the projects provide “the moment of transformation that one sees from a passive receiver of knowledge to a producer of knowledge.”

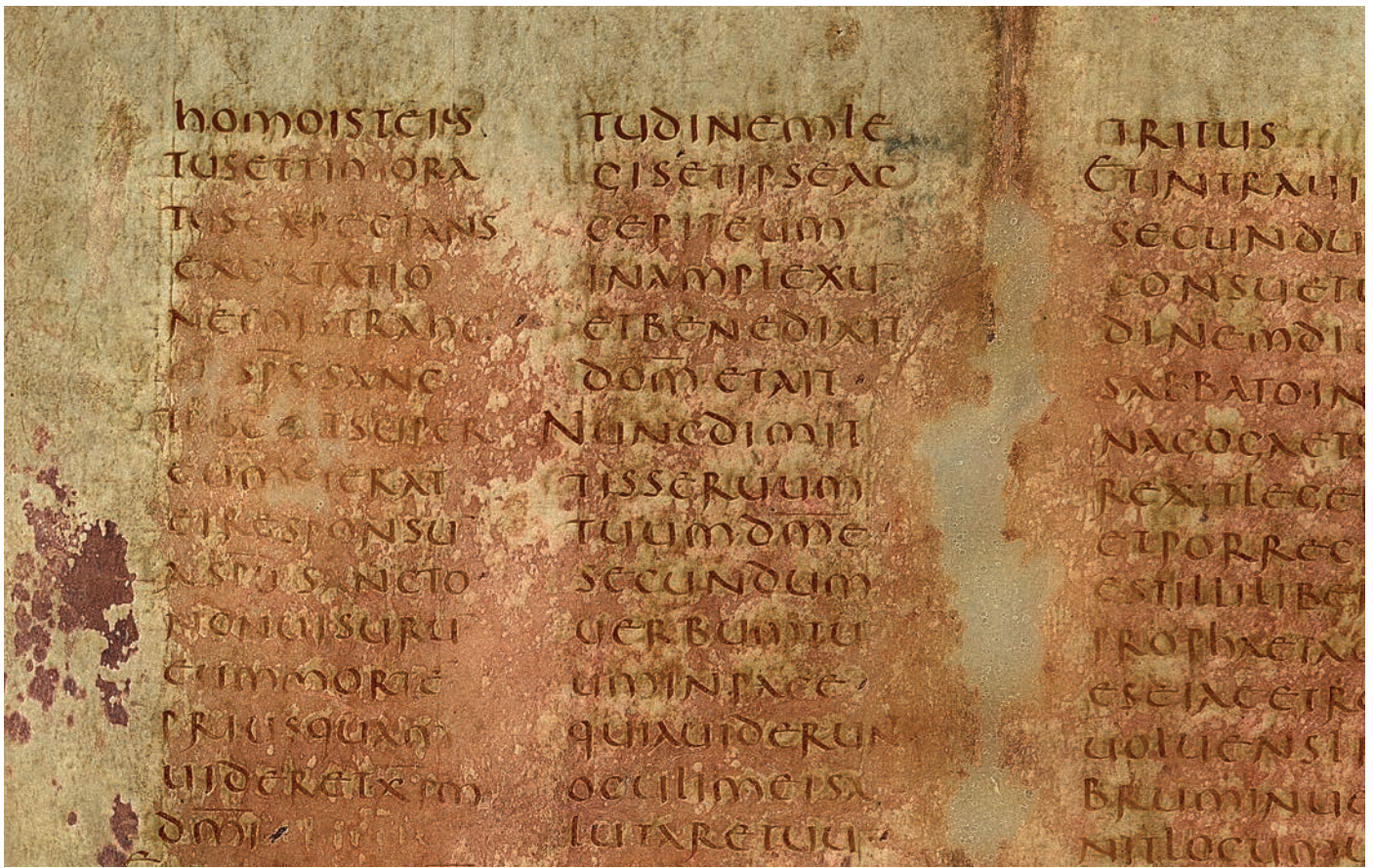
**H**e welcomes the interest of anyone whose enthusiasm for the quest might begin to match his own. He's already working with students from computer scientist Henry Kautz's lab to see how machine learning—in the form of computer-aided optical character recognition—could be used one day to read rediscovered texts, moving through manuscripts more efficiently than individual scholars can. Students of optics professor



**WORLD TRAVELS:** Heyworth's group, the Lazarus Project, crosses the globe to image imperiled cultural heritage objects. During a 2015 visit to the country of Georgia (top), Heyworth examined a charred manuscript in the Georgian National Archives (above).

Wayne Knox may collaborate with Heyworth to determine how the University's collection of dime novels—one of the world's largest—can be accessed, given that many of the books can be opened only a few degrees.

And in a freewheeling assemblage almost unheard of in the fairly solitary world of humanities research, Heyworth spoke at Rush Rhees Library one afternoon last autumn to a room filled with students, faculty, and librarians from the University and RIT—not hand-picked people but anyone interested in attending. He explained dimensions of the projects in which they could involve themselves. For example, in the late 15th century, scholars started using chemical reagents—chemicals used to “revitalize” texts for 10 seconds or a minute, after which time the chemicals would permanently stain the pages, obscuring the text even further. “Anyone with a background in this?” Heyworth called out, on the hunt for



**BEFORE AND AFTER:** A page from the Codex Vercellensis—thought to be the earliest manuscript of the Old Latin Gospels—before Heyworth and his team complete multispectral imaging and processing (above) and after (top).

chemists. “I have recipes for reagents and medieval inks and would love to do some tests in a lab, to see if we can find ways to see through reagents.”

“We’re going to attract people from communities that don’t talk to each other,” says Easton, who is still amazed that his scientific work now brings him into regular contact with some of the world’s most eminent humanistic scholars. “We don’t speak the same language. So that’s going to be a challenge,” he says of the team’s plans.

But Heyworth has no doubts. “I think in this new digital world, it’s absolutely necessary to integrate the sciences with the humanities,” he says. “Without that we can’t have true innovation. We can’t move forward. I think we owe it to our students and to our next generation to train them in ways to work together, to do just this.”

And Phelps is convinced that Heyworth and Easton’s proximity and the chance to bring together experts at Rochester and RIT in the service of preserving world cultural heritage will bring as-yet unimagined breakthroughs.

In humanities research that relies on technology, “we’re always using derivative technologies. They were made for some other purpose, and we try to use them for humanities work,” he says. “Greg’s move to Rochester, with the collaboration developing with RIT, brings a real opportunity for humanities people and the unique problems we face in preserving and restoring treasures of our culture. It’s an opportunity to have the problems of our fields in the humanities—the problems of historical research—drive technological development.”

Whether it’s ancient manuscripts or 20th-century political ephemera, the world is filled with objects waiting to be read—to be rescued, before they crumble into dust, deteriorated beyond saving. And everyone has a stake in their preservation, says Phelps.

“I don’t see these projects as ivory-tower scholarly projects,” he says. “Of course, they have a scholarly component, but this is the heritage of us all.”

Says Heyworth: “We have an obligation not only to try to make this technology available for free or for almost no cost—but also to train students to do this, to expand our capacity.”

Cultural heritage imaging is only in its infancy, and there is much to explore, he says. He calls what lies ahead “the future of the past”—a bold and unfamiliar future.

“I predict that in the next 10 years,” he says, “textual science is going to revolutionize the kind of work that’s been going on in the humanities.” 