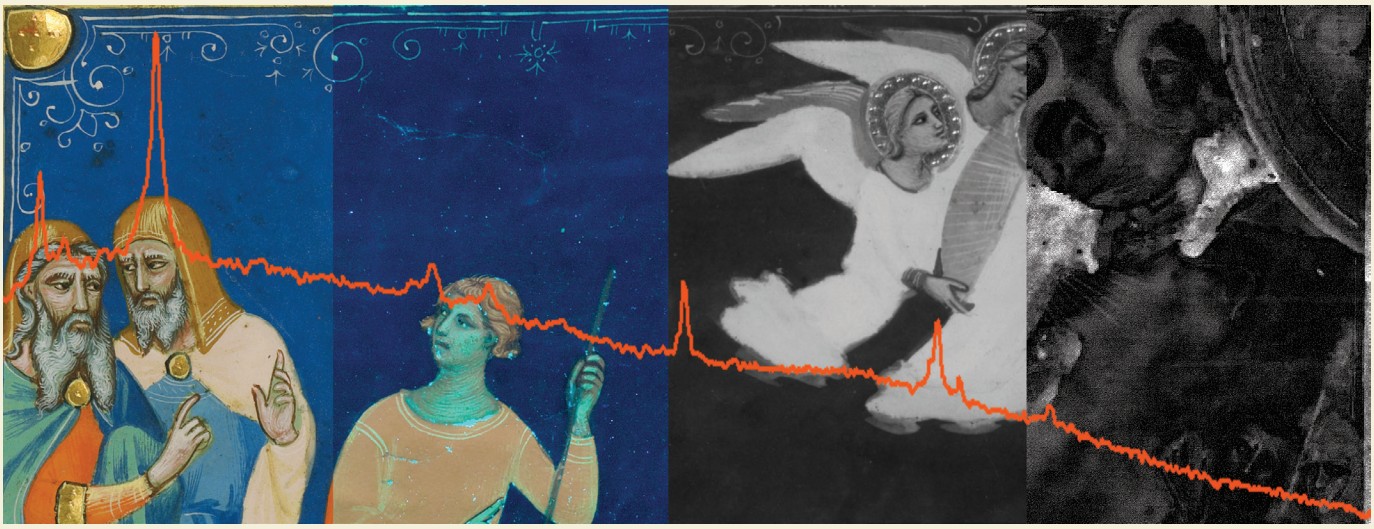
DISCO
DISCO
DISCO
DISCO

At the Getty Conservation Institute, it's much more than a dance craze



Composite visualization of four images (visible, UV, infrared, and x-ray images) and spectroscopic data from adjacent areas of a single fourteenth-century manuscript leaf (JPGM collection, Ms. 80b, verso)

The disco ball—the iconic symbol of dance and music popular in the 1970s—with its thousands of micro-facets reflecting light in a multitude of directions, may be seen as a metaphor for the sometimes dizzying amount of information generated by conservators and conservation scientists. As it becomes ever more abundant, how to fully synthesize this data to produce new insight is now the challenge.

Researchers in the field of cultural heritage, as in other disciplines and areas of research, are creating an enormous amount of data gathered during an analysis or treatment campaign for a work of art or cultural heritage site. The same is true for curators and art historians, who increasingly incorporate information from technical studies into their work. Taken together, the different types of data these researchers collect—images, text, and analytical measurements—inform the way conservation treatment programs are undertaken and how historic, artistic, technological, and cultural interpretations are made.

Through the DISCO (Data Integration for Conservation Science) project, which will develop computer-assisted data integration tools that will facilitate extraction and sharing of new information by a broad community of users, the Getty Conservation Institute (GCI) is working to enhance the ways scientific and technical studies contribute to the conservation and understanding of works of art.

A major challenge for researchers is the sheer volume and diversity of information attainable. As instrumentation capabilities continue to advance, a single researcher's ability to collect data from a variety of sources is beginning to surpass his or her ability to manually fully analyze, interpret, and synthesize the information generated. Even simply comparing data from a single object can be a challenge if they are in different formats; as a result, all the data relating to an object (or group of objects) are rarely examined together. Consequently, subtle, but important phenomena or relationships can be difficult to identify in an overwhelming data stream.

Researchers need a way to integrate all the data relating to the topic under study—whether that be a single object, group of objects, artist, or artist's workshop—into a concise and accurate representation that facilitates making comparisons, finding correlations, and ultimately leading to new insights and discoveries.

"An ideal situation would be one in which we could collate diverse data sets, and look for relationships that we didn't *a priori* know existed," said Catherine Patterson, associate scientist at the GCI. "Such an ability would have been transformative in our project with the J. Paul Getty Museum to use the materiality of manuscript illuminations and panel paintings to try to better understand fourteenth-century Florentine workshop practice."

The need for data integration is not unique to cultural heritage. Other data-rich disciplines, such as astronomy, medicine, chemistry, finance, and gaming have begun to develop resources attuned to the needs of their respective stakeholders. Individual aspects of data integration for cultural heritage, including the building of research networks to exchange data or the visualization of conservation technical images, have also begun to be addressed in recent years. But the overarching challenge of how to integrate all of the different types of data relating to cultural heritage research has not yet been addressed. Until now.

In 2013 the GCI convened a meeting of thirty experts to address this need. Made possible by the generous support of Dan Greenberg and Susan Steinhauser, the meeting explored the ways computer-assisted technologies could help cultural heritage researchers integrate different types of data—including those from different researchers and different institutions—in a way that facilitates the extraction,

Summary visualizations of research data from a fourteenth-century Italian manuscript leaf in the JPGM collection (Ms. 80b, verso)

sharing, and understanding of new information by a broad community of users.

The experts were gathered from a wide variety of fields: conservation, conservation science, art history, imaging science, data visualization, data and information science, astronomy, computer science, medicine, and software development. During two-and-a-half days of animated meetings, the group discussed the state of the field in both cultural heritage research practice and computer-assisted technologies, identified new avenues of research that would be made possible or enhanced by data integration, suggested measures for implementing and supporting computer-assisted solutions to data integration, and identified priorities for action.

"After the experts meeting in 2013, we realized that a new paradigm needs to be created for the way data are collected, stored, accessed, shared, analyzed, and interpreted," said Karen Trentelman, principal investigator for the DISCO project and senior scientist at the GCI. "Through DISCO, we hope to enable researchers to more effectively leverage scientific and technical studies to advance cultural heritage research."

It is in part from the recommendations of that 2013 meeting that the GCI formally developed DISCO. The project, with support from the Seaver Institute, has since conducted additional research, including a use-case analysis by interviewing individuals at a wide variety of cultural heritage and research organizations around the world to better determine how to positively affect the state of data integration within conservation science. Surprisingly, the findings showed that often organizations are not able to take full advantage of, or contribute to, many new digital humanities projects because they are still working on foundational tasks such as organizing and managing their own data in a way that facilitates data integration.

As a crucial first step, DISCO is working to meet this most urgent need in the field. In order to make the biggest impact, the project is currently focused on creating open source tools that allow researchers to create and manage digital scientific data that will be organized according to a uniform standard.

"The GCI, with its resources and expertise, is uniquely positioned to undertake a foundational project like DISCO, which has the opportunity to help redefine how conservation scientists manage their data, and thus, improve the field as a whole. Once the groundwork is laid, true data integration becomes more feasible and something that we can tackle," said Alison Dalgity, senior project manager at the GCI.

Ultimately, multiple types of data, from both individual researchers and groups of researchers at different institutions, will be connected through linked data. This linked data will facilitate interrogation, visualization, and data interpretation, helping researchers, for example, draw comparisons and correlations between different works of art, different studies, and different points in an object's history. Linking data from multiple sources will add value to each individual data source, and the community of experts built around the shared data will improve the quality of data interpretation.