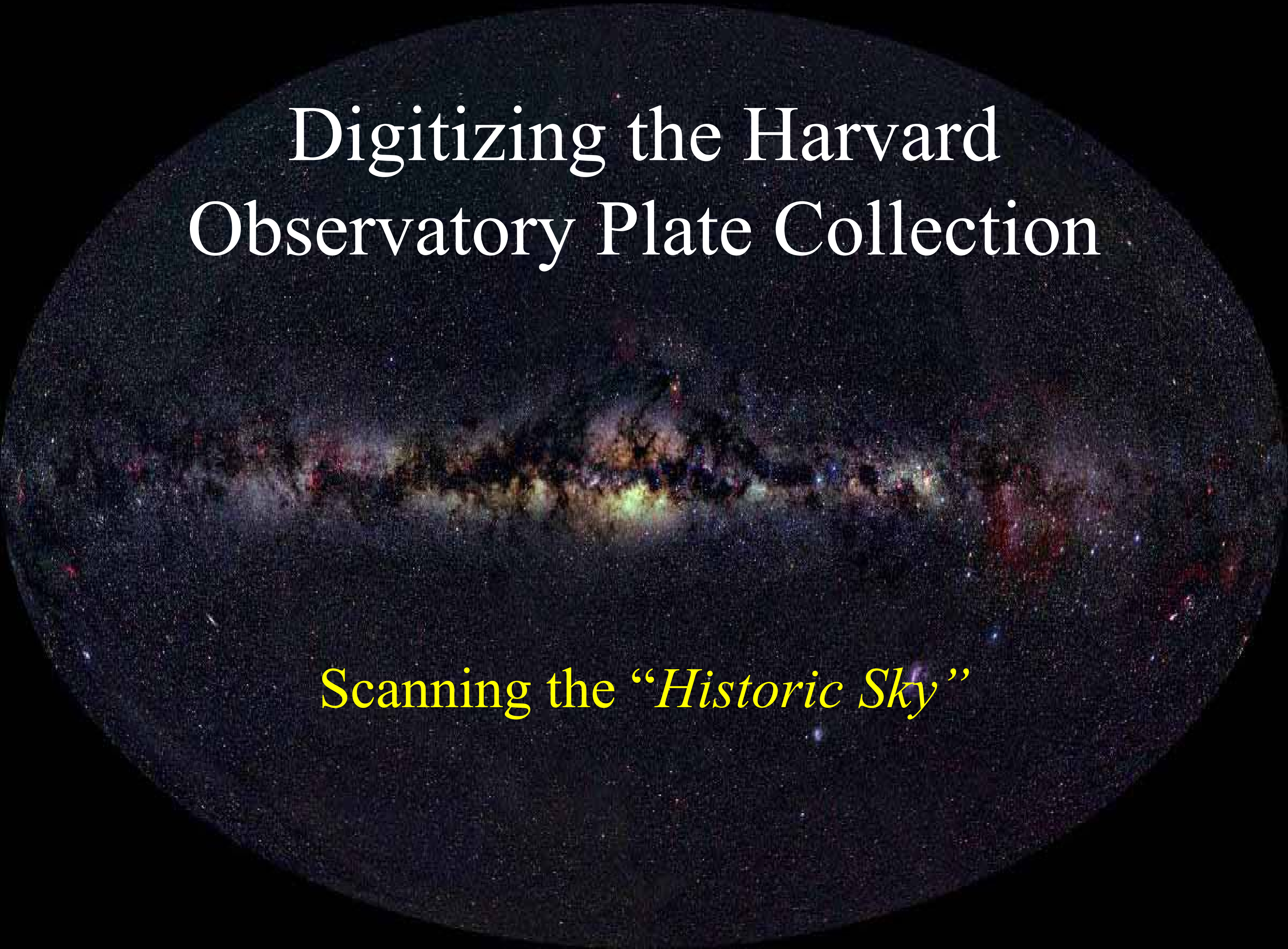
Digitizing the Harvard Observatory Plate Collection

Scanning the "*Historic Sky*"



Our Goals: Find Funding to Construct a Scanner and Digitize the Harvard Astronomical Photographic Plate Collection. Make the results available in Online Storage.

Jonathan E. Grindlay – Harvard Professor of Astronomy

Elizabeth Griffin – WG Chair IAU Digitization and Preservation

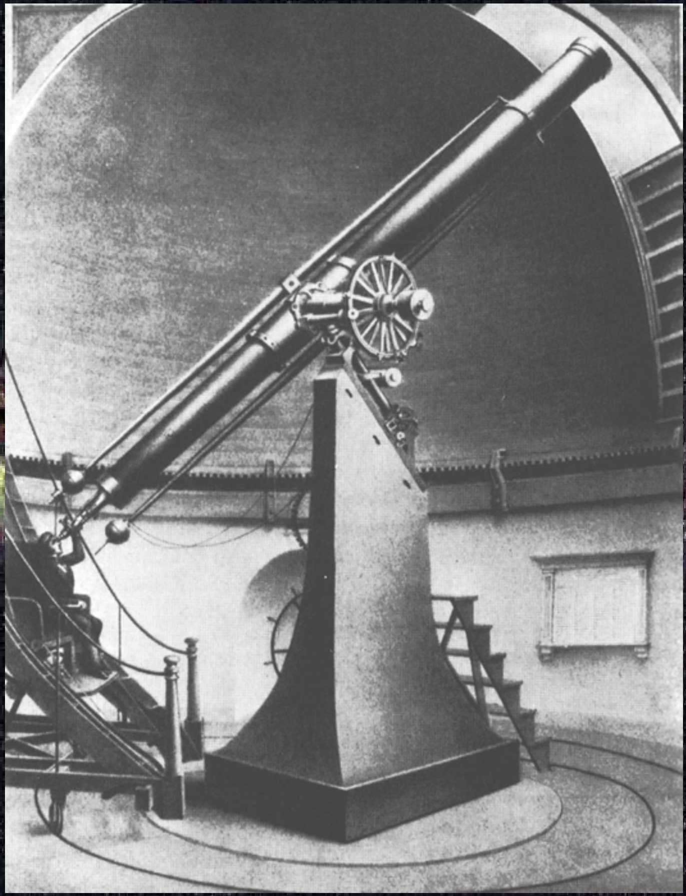
Alison Doane – Acting Curator of the Harvard Plate Stack

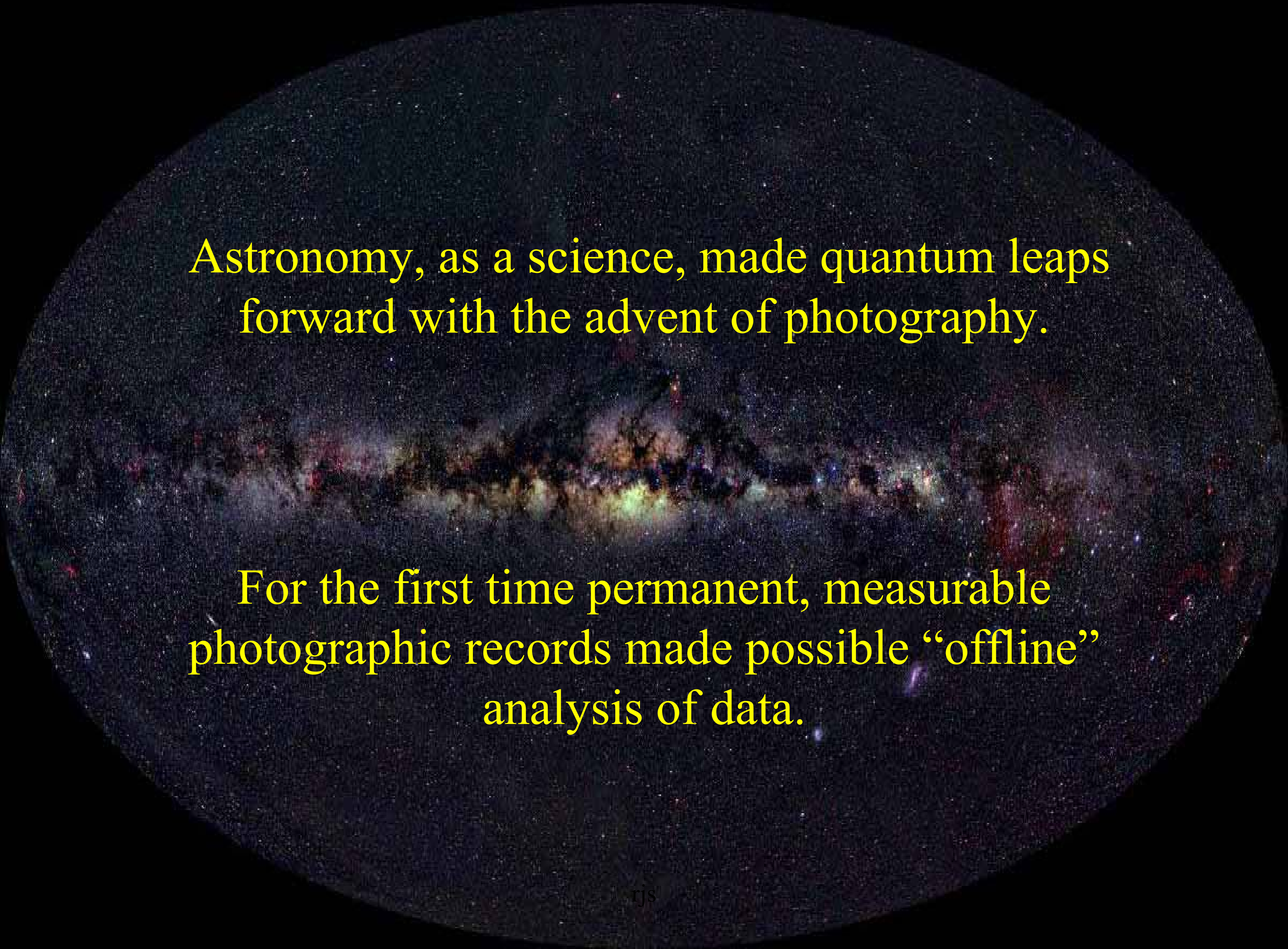
Douglas J. Mink - Software and Data Archivist

Bob Simcoe – Volunteer Associate & System designer

Before photography, astronomers' eyes were their only sensing device and hand drawing was the means of permanent recording.

This severely limited the science they could accomplish.

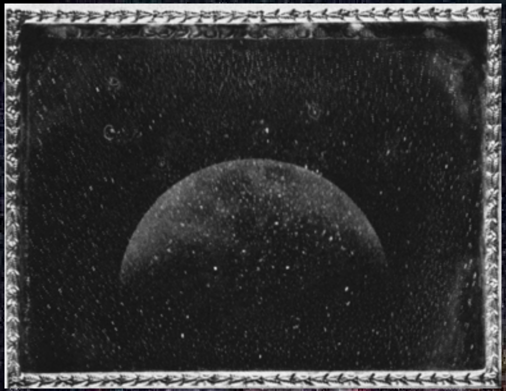




Astronomy, as a science, made quantum leaps forward with the advent of photography.

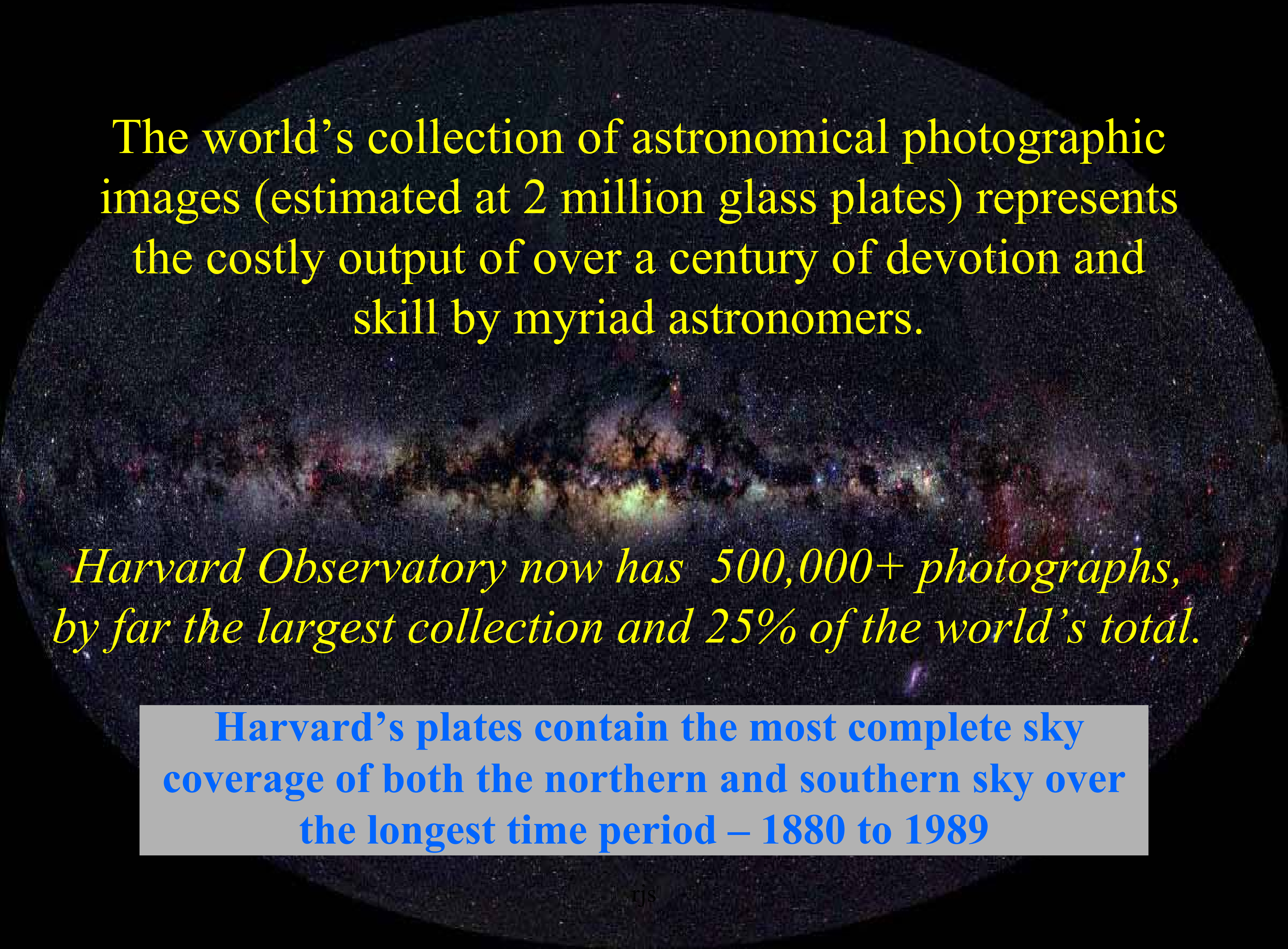
For the first time permanent, measurable photographic records made possible "offline" analysis of data.

The first daguerreotype of the moon was made by American physiologist J.W. Draper in 1840, involving a full 20 minute exposure.



The first star was not recorded until 1850, when director of Harvard Observatory, W.C. Bond and Boston photographer J.A. Whipple, took a daguerreotype of Vega.

The first photographic sky surveys were done at Harvard during the period of 1882-1886. Each photograph covered 15 degree squares of sky and recorded stars as faint as 8th magnitude.



The world's collection of astronomical photographic images (estimated at 2 million glass plates) represents the costly output of over a century of devotion and skill by myriad astronomers.

Harvard Observatory now has 500,000+ photographs, by far the largest collection and 25% of the world's total.

Harvard's plates contain the most complete sky coverage of both the northern and southern sky over the longest time period – 1880 to 1989

Since the 1980's, astronomers have largely abandoned the use of photography.

They now use CCD imagers that directly digitize sky images at the telescope.

Harvard's collection, and other collections around the world, are already nominally in the public domain. But as a resource they are seriously under-exploited. The main reasons are:

(a) **Lack of catalog information in digital form about the plates**

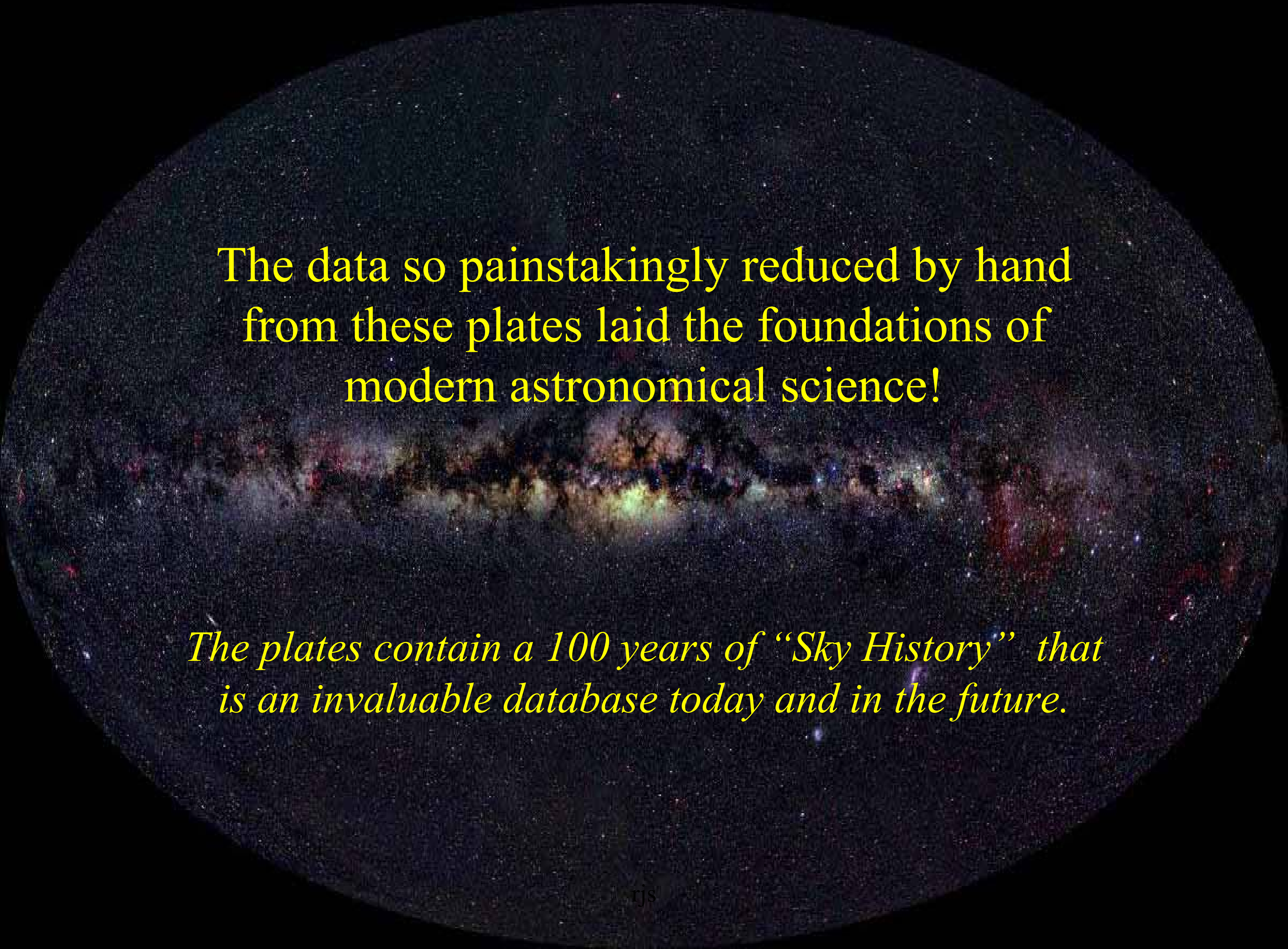
(b) **Lack of digitized data from the plates**

The modern tools of astronomy require digital data!

Harvard's and other plate libraries are in danger of being lost because of low use.

It typically requires travel, physical handling of the plates, and use of antiquated equipment to study the plate contents.

The low use puts pressure on the collections' budgets for storage space, personnel, and maintenance of very old equipment.



The data so painstakingly reduced by hand from these plates laid the foundations of modern astronomical science!

The plates contain a 100 years of "Sky History" that is an invaluable database today and in the future.

The “analog” storage of data on Harvard’s photographic plates represents about 200 Terabytes of digital data.

Until now, despite the desire to convert the plates to a digital form, it was not technically or economically feasible to either scan the plates or store the resulting data online.

Now both are possible!

The last plate scanning machine designed in the US (the PMM, ~1988) scanned ~ 20,000 plates over it's lifetime. It is currently retired. No machine of this type has been designed in the United States since then!

Machines of the PMM era took ~1- 4 hours to scan a

single 14 x 14 inch plate.

To scan Harvard’s library of historic plates in a 3-5 year timeframe, we need a machine that can scan ~ 200 times faster than the machines designed 20+ years ago.

To meet astrometric, photometric, and archival goals, the machine needs sub-micron positional accuracy, at least 12 bits photometric density range, and a scan speed that allows human handling to limit the “average plate processing” time.

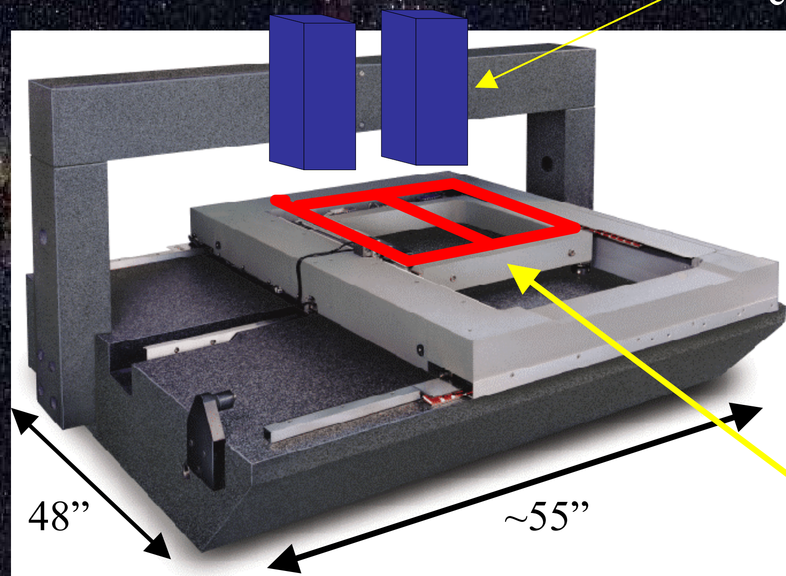
Using technology common to semiconductor wafer and flat panel display inspection stations, a machine can be built today that can do ultra fast, ultra precise scanning.

The scanner will scan an 8 x 10 inch plate in about 20 seconds. It will scan a 14 x 17 inch plate in a little over a minute, generating enough data in that time to fill a DVD.

The scanner must provide astrometric and photometric accuracy while generating archival quality digital data.

Two cameras can scan two 8 x 10 inch plates simultaneously or one camera can scan a single 14 x 17 inch plate

Cameras are capable of ~10 frames per second.



Each camera takes ~ 250 pictures in 50 seconds (4 exposures for each of ~63 sites for an 8 x 10 plate).

Special Fixturing to hold 14 x 17, 14 x 14,10 x 10, 2(8 x 10), and smaller plates as needed.

Online storage is the only practical – and the most useful – way to store this deluge of data.

Hard disk storage has become more cost effective than photographic film or paper.

It is now less than $1 per Gigabyte.

The “Virtual Observatory”- envisioned for archiving current and future observational data- should have as a foundation the data of the “Historic Sky”.

Much of what is on these plates has not been mined, because most previous data reduction was done by hand with simple measuring machines.

The worldwide astronomical community naturally looks to Harvard to lead the way in the effort to make 100+ years of collected “Historic Sky” available in digital format and online.

Digitizing the plates constructs a “Telescope into the Past” for the “Virtual Observatory”.

The foundation and development of Harvard Observatory has been due in large part to private gifts. Harvard's rich heritage of photographs is largely due to the generosity of Miss Catherine Wolfe Bruce, Mrs. Henry Draper and the hard work of women like Williamina Fleming, Antonia Maury, Henrietta Leavitt, and Annie Cannon.

It would be very fitting to have private gifts or grants preserve and make this "Historic Sky" treasure trove accessible to future generations of astronomers.

Help us put the resources together to save this treasure and make it available to the scientific community in the form that is required today!

What will it take to digitize Harvard's plate collection and store the results online?

Total time - 5 years

Total Program - $5M

Total storage ~ ½ Petabyte

Run rate ~ $1M/year

Year 1 – Develop and test the scanner

Year 2, 3, 4 – Scan in the bulk of the plates

Year 5 – Set up for scan on demand, finish archival storage, document the project

Storage will be purchased incrementally throughout the project

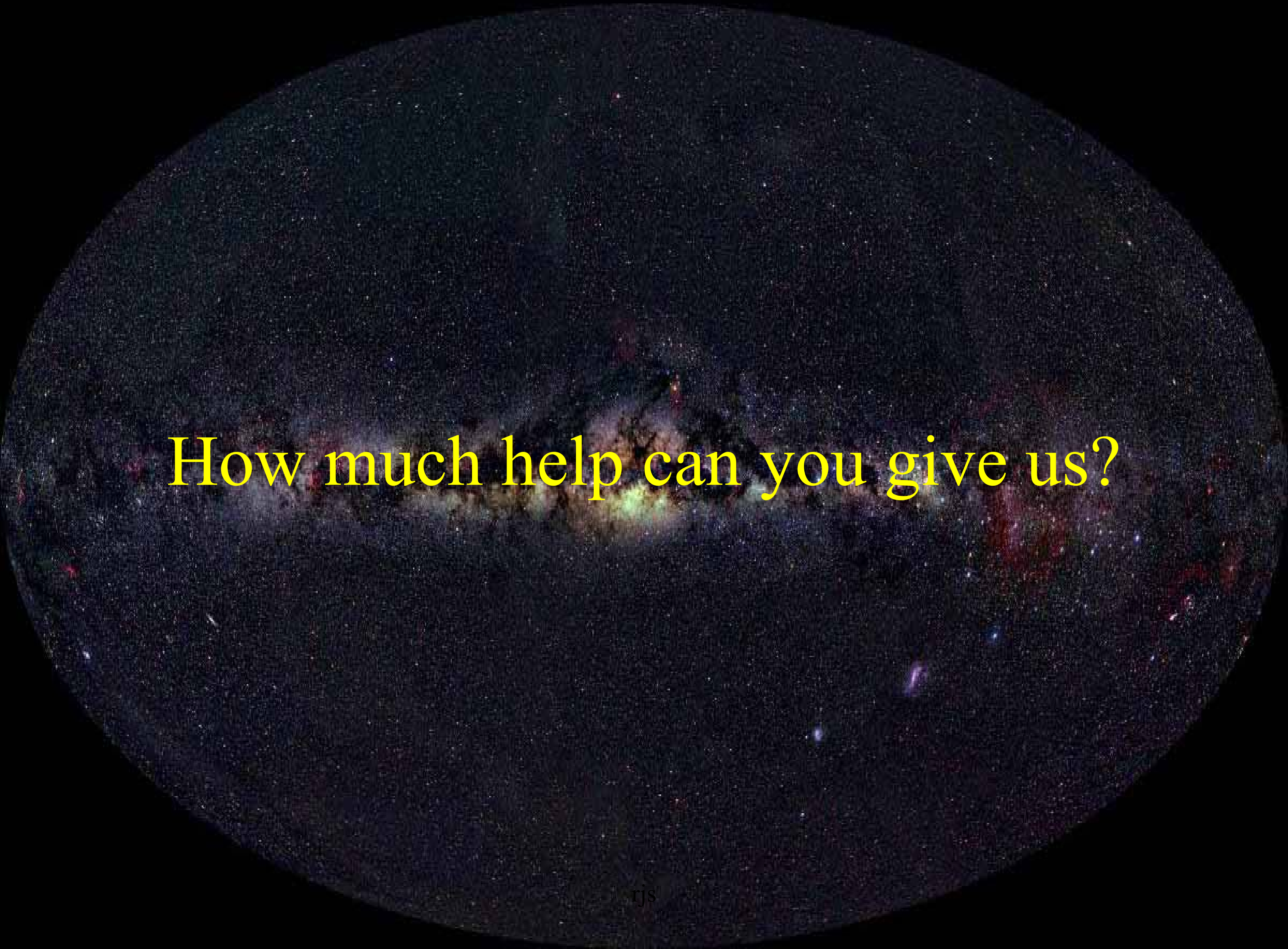
The scanner and facility development ~$250K

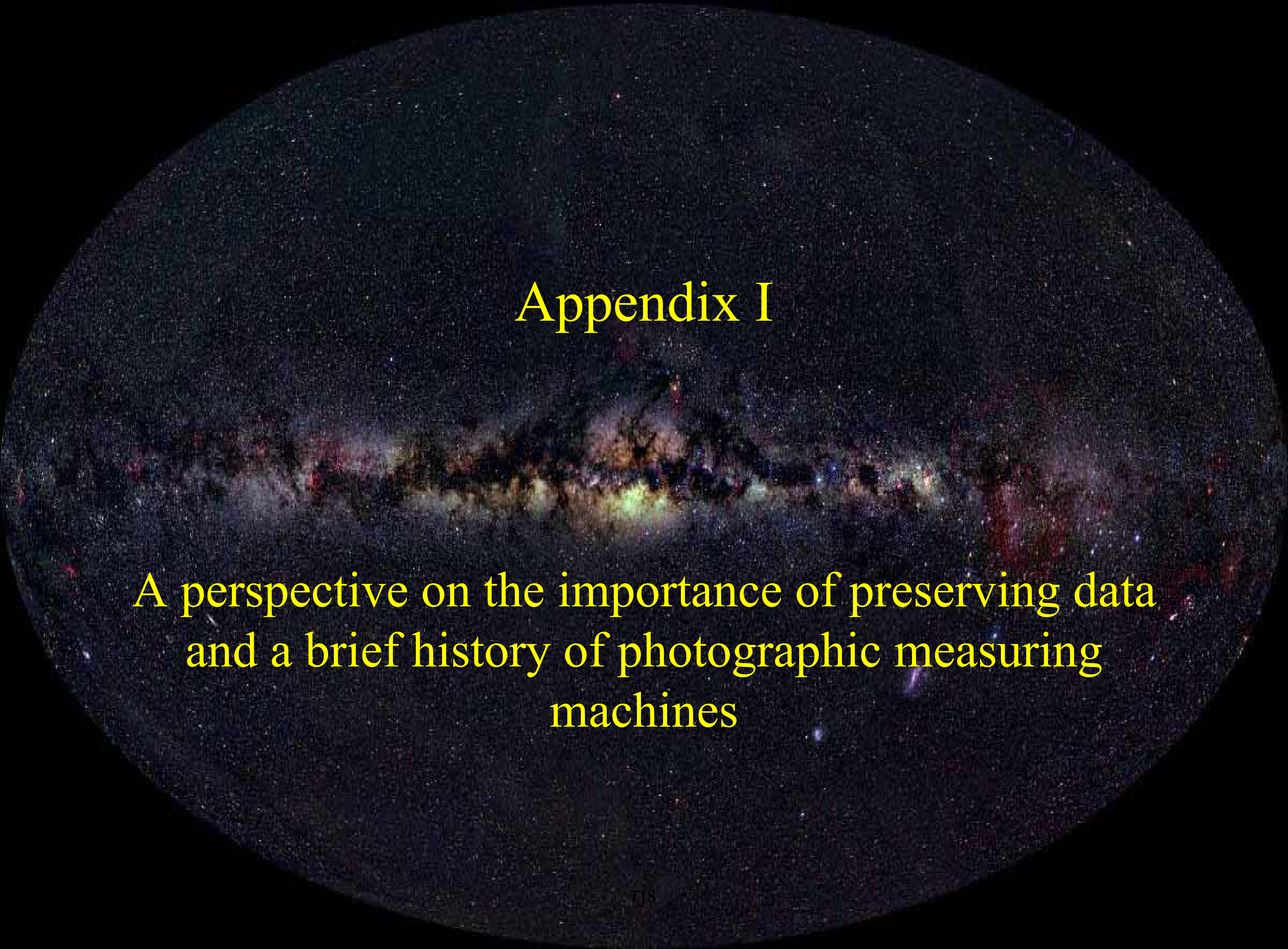
People costs to scan the ~ 500,000 plates

4 plate handlers and 2 programmers ~$2.5M

The 500 Terabyes of reliable online Storage

+ $ to endow its operation ~ $2.25M


How much help can you give us?
rjs



Appendix I

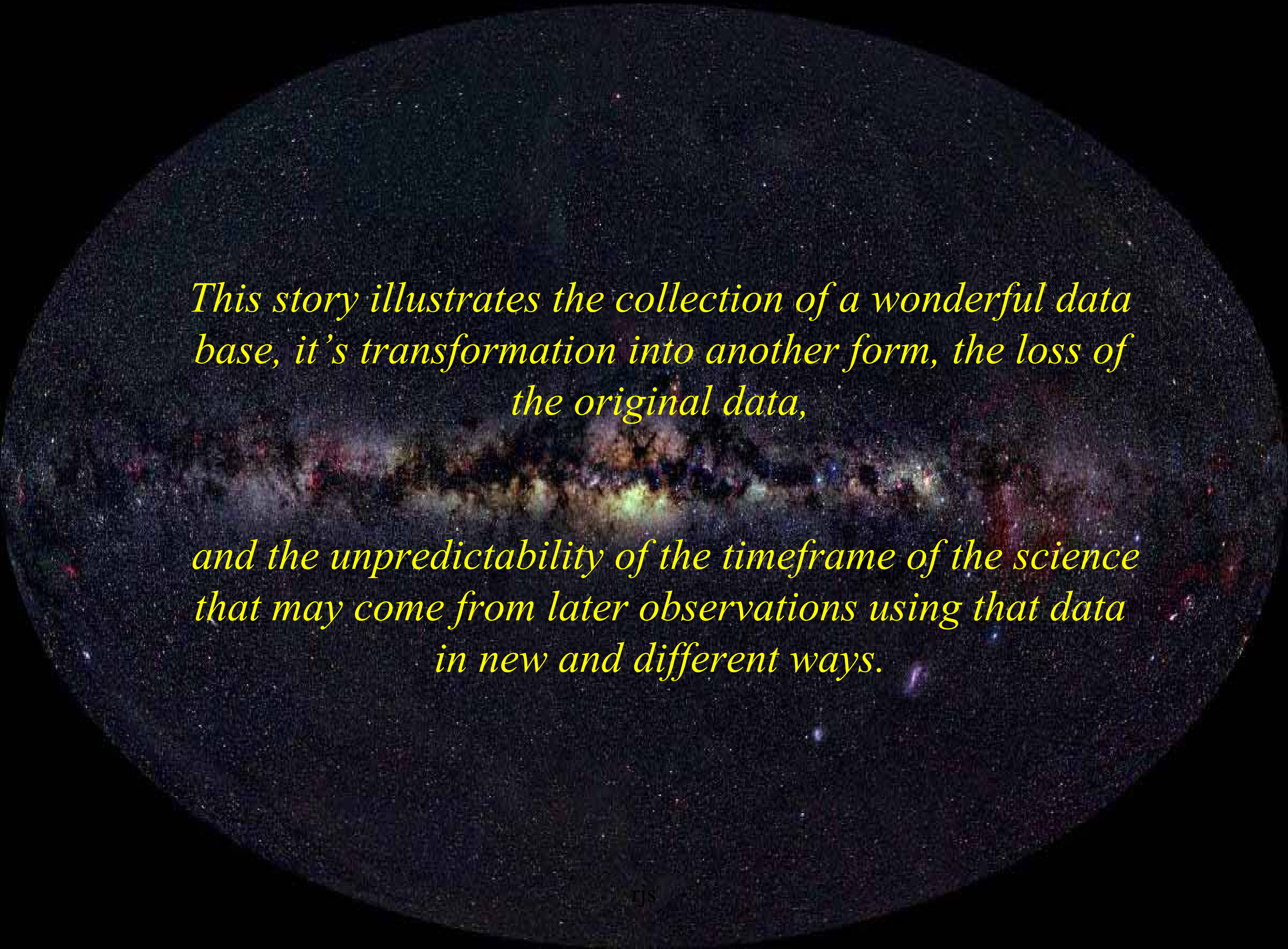
A perspective on the importance of preserving data and a brief history of photographic measuring machines

"The first optical sky survey which we know about is a poem. *Phaenomena*, by Aratus (c. 350 BC). It set out the shape of the constellation figures and the stars which form them. Aratus' survey is based on one now lost, by Eudoxus, who visited Alexandria for the

purpose of examining the records of observations kept in the library and for consultation with the scholars there – these observations were originally made about 2000 BC, perhaps from Minos in Crete.

Aratus gave some interpretations from his survey in the form of navigational and weather lore; a more scientifically valuable interpretation was made by Hipparchus two centuries afterwards. He noticed that parts of the sky described by Aratus (the southern parts of the constellation Argo) were no longer visible above the horizon seen from Greece, and other parts that were visible had no described constellations. This was due to a change in the tilt in the earths axis relative to the stars – Hipparchus thus discovered precession.”

P.G. Murdin – Welcoming address: Two Thousand Years of Optical Sky Surveys
Proceedings of the conference on “Digitised Optical Sky Surveys”
Edinburgh, Scotland, June 1991

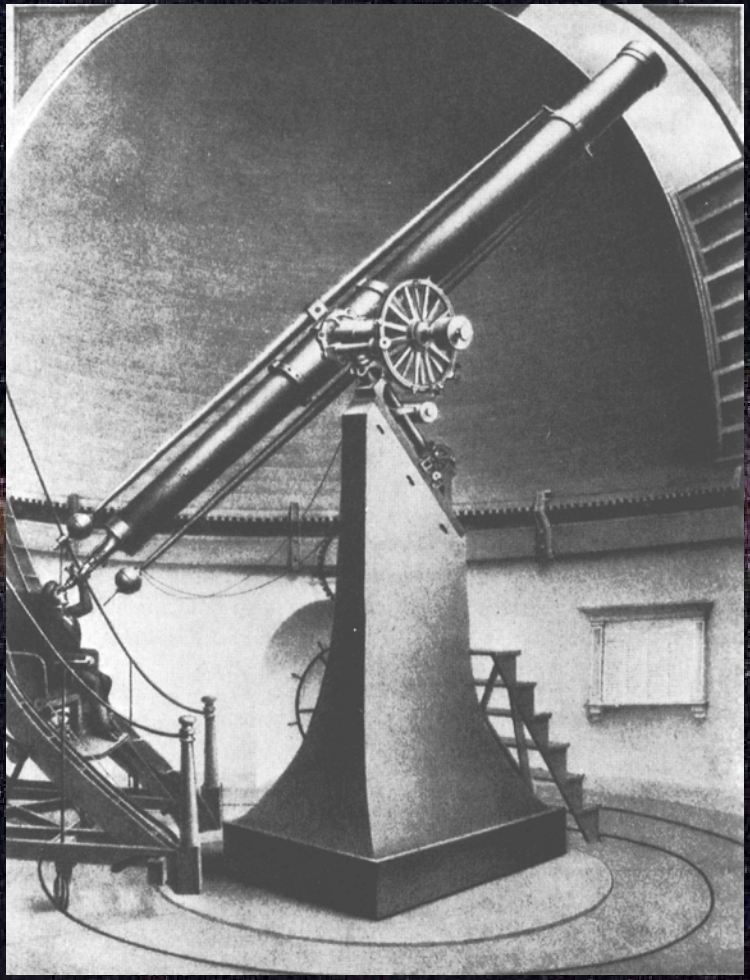

This story illustrates the collection of a wonderful data base, it’s transformation into another form, the loss of the original data,
and the unpredictability of the timeframe of the science that may come from later observations using that data in new and different ways.
rjs

“The earliest record to give *quantitive* information on the brightness of stars is in Books VII and VIII of the *Almagest* by Claudius Ptolemaeus (Ptolemy) (c. AD 100 – c. AD170) The star catalogue of the *Algamest* lists ~1000 stars in 48 constellations and gives each a magnitude on a scale from one to six. Ptolemy says almost nothing in the *Almagest* about how he defines magnitude. The original work is no longer available but numerous manuscripts from the ninth to sixteenth centuries have survived.”

“The measurement of starlight” J B Hernshaw

Before photography, astronomers’ eyes were their only sensing device and hand drawing was the means of permanent recording.

This severely limited the science they could accomplish.



Photography enabled “offline” instruments to be used to measure the sky. There are three fundamental types of measurements.

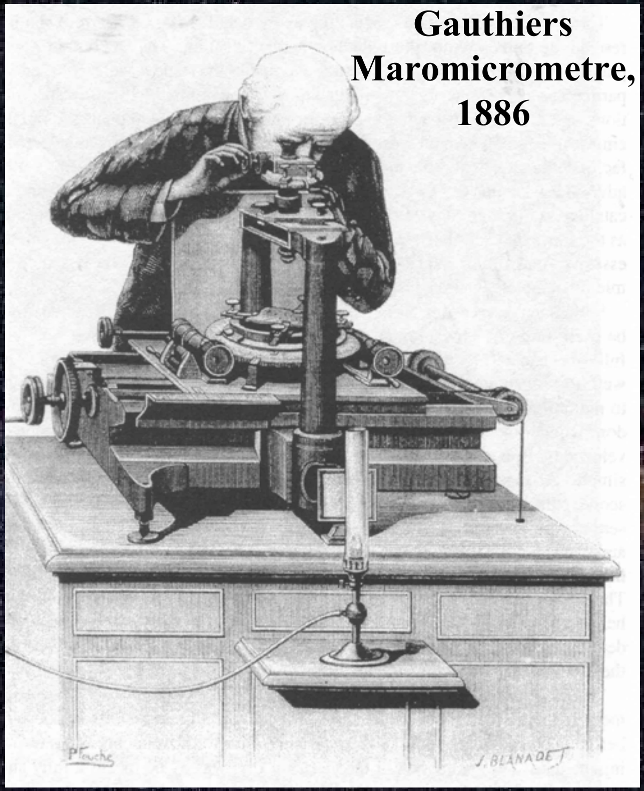
Astrometric –the proper position and motion of stars

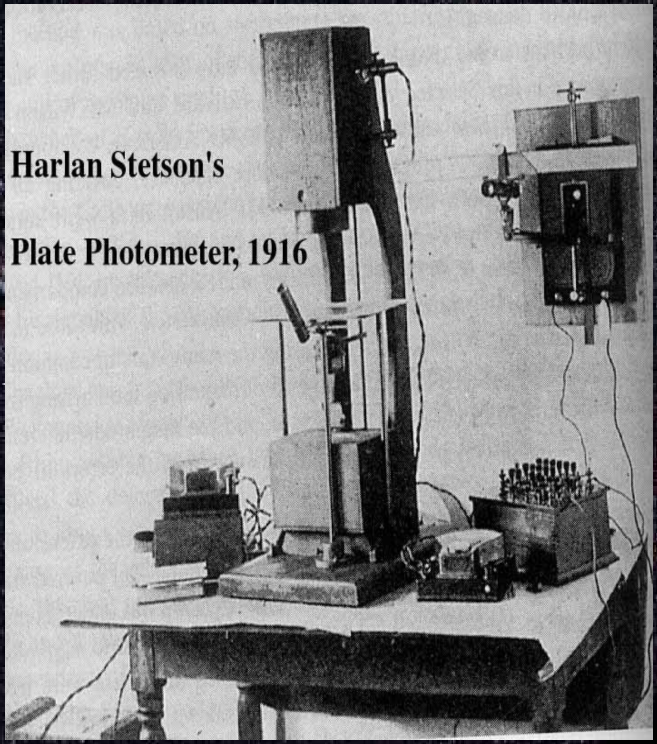
Photometric –the brightness of stars

Spectroscopic –the positions of the absorption lines of a star’s spectra.

Photo –metrics

Measuring brightness





Astro – metrics

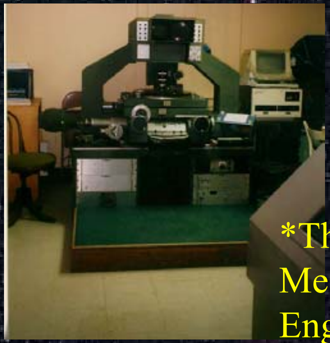
Measuring positions

Harvard's first astronomy "computers"

Automating the Measurements

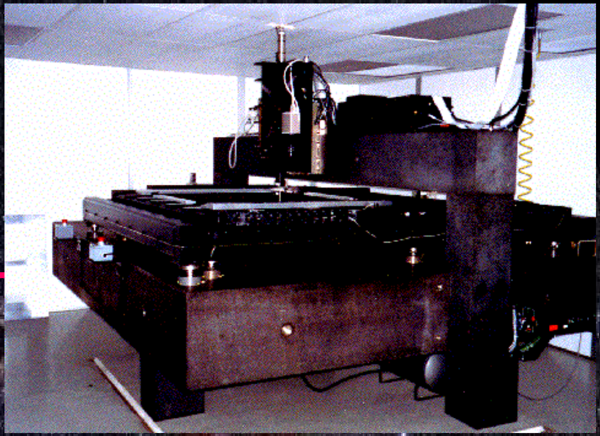
1886

*Gaertner single screw engine 1916



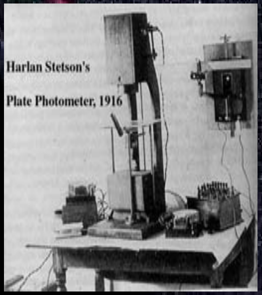
*The Grant 2 Measuring Engine-1967

**PMM (NRO -1988)

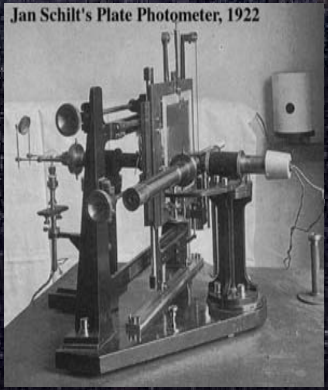


Astrometric

Photometric



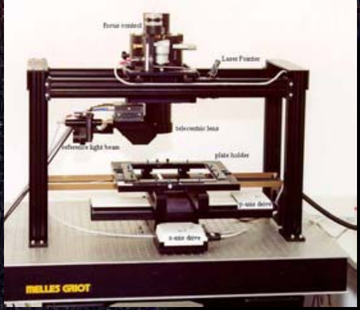
*1916



Jan Schilt Photometer - 1922

*Perkin-Elmer

PDS -1980



Tautenburg-1995

*http://www.astro.virginia.edu/~rjp0i/museum.html

**http://www.nofs.navy.mil/projects/pmm/pmm_caption.html

Work that took many “women years” of labor at the end of the 19th century can now be done by scanning machines and computers in seconds or milli-seconds.

Appendix II

Another significant database digitization opportunity-

worldwide in scope

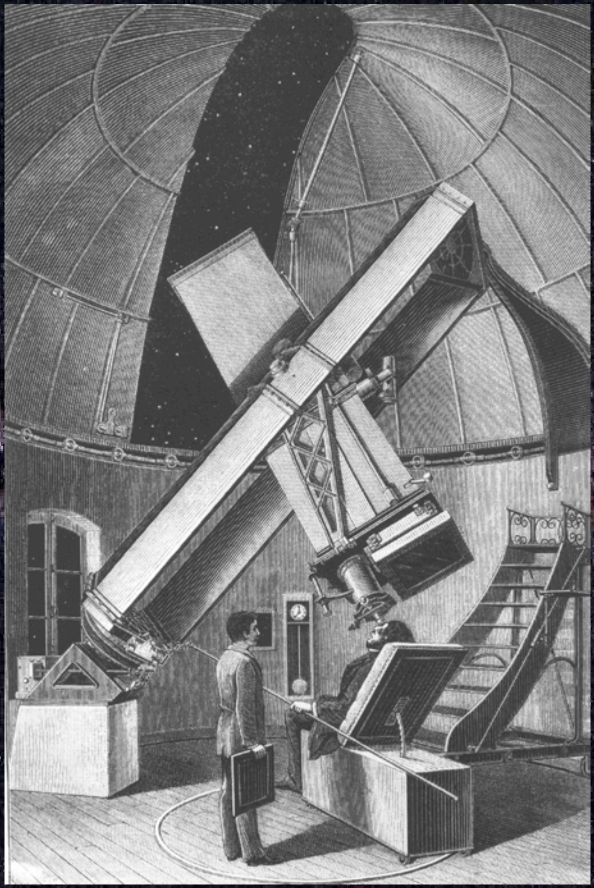
The Astrophotographic Catalog (AC) and La Carte du Ciel

La Carte du Ciel was the first large international science effort. It began in 1887 with 18-20 participating observatories around the world. The goal for the AC photos was a ten minute exposure to determine accurate

positions for all stars brighter than 11^{th} magnitude. Each observatory was assigned a specific zone between two parallels of declination. It took until the 1930’s to accumulate and *hand measure* 22, 660 plates for the AC.

Each participating observatory used a telescope with a 16cm lens, photographing overlapping sections of the sky.

This work did not officially end until 1962. Data from the AC has been updated as late as 1998.



La Carte du Ciel

A second set of 30 minute exposures were taken for the "Carte du Ciel" which was designed to capture the sky to at least magnitude 14. This work proved to be too expensive to complete and many participating observatories had to abandon the job of taking the plates and measuring them.

There remains from this effort about 20,000 plates that have never been measured, capturing most, but not all, of the sky.

Digitizing the plates of "La Carte du Ciel" would help preserve and make useful this "labor for science" for current and future generations of astronomers.