



Australian Government

Geoscience Australia

A map of Australia with a colorful, textured overlay representing remote sensing data. The colors range from green and yellow to blue and purple, indicating different land use or environmental data. The text "Australia's Remotely Sensed Data Archive: The Next 25 Years" is overlaid in red.

Australia's Remotely Sensed Data Archive: The Next 25 Years

Stuart Barr

Geospatial and Earth Monitoring Division

Introduction

- **The Australian Centre for Remote Sensing (ACRES) is Australia's major public satellite remote sensing organisation**
- **It was established as the Australian Landsat Station in 1979 and has operated continuously ever since**



What data are collected?

- ACRES collects and archives remotely sensed data of Australia, New Zealand and Oceania
- Principal holding is of Landsat data
- Data sources include:
 - Landsat 2, 3, 4, 5 and 7
 - Radarsat-1
 - ERS 1 and 2
 - MODIS (Terra and Aqua)
 - ASTER
 - ALOS
 - SPOT 1, 2 and 4
 - JERS-1
 - NOAA - AVHRR



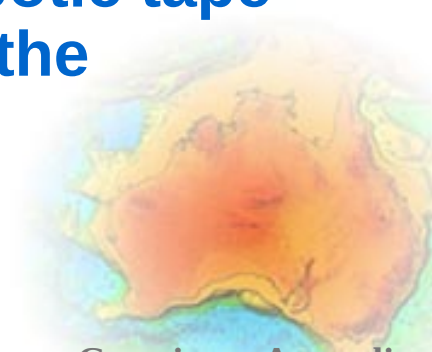
Uses of the data

- **Data are used for a wide range of purposes including**
 - **Agriculture for crop health, acreage and yield estimation**
 - **Australian Greenhouse Office for carbon accounting**
 - **State Government monitoring of land clearing**
 - **Mining prospectivity**
 - **Map production**
 - **Weather prediction**
 - **Fire detection and fire scar mapping**
 - **Climate change research**



Data characteristics

- **Imagery data sets are large!**
 - A ten-minute satellite overpass can contribute 10 GB
 - Total unique data holding exceeds 200 TB
 - Limited opportunity for compression both by reason of data structure and resistance by the user community
- **Data are stored on tape (presently SDLT)**
 - A significant percentage is stored in robotic tape libraries, the balance in cabinets within the computer room



Data characteristics

- **Data sets are added to the archive automatically**
 - Metadata creation is also automatic
 - Ancillary data necessary for processing are also captured automatically
- **Data management performed by an in-house developed archive server that tracks the location of all data sets**



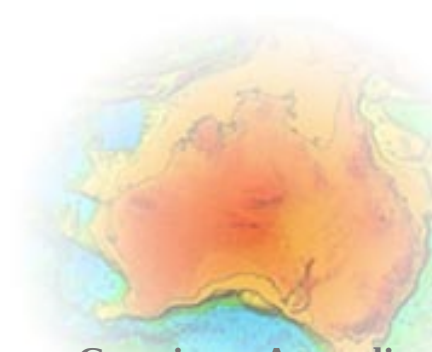
Data characteristics

- **Data are mostly stored in raw form and processed to imagery on demand**
 - This enables the latest improvements in processing algorithms to be incorporated
 - An in-house developed archive format enables recovery of original data
 - Significant ancillary files are also required to permit processing
- **Processing software is highly specialised**
 - The processing software retains compatibility with “old” data sets



Trustworthiness

- Users expect and demand that the raw data are maintained without corruption or distortion, however . . .
- The data are processed prior to use and the processing algorithms are updated periodically, but . . .
- The provenance of the processing algorithms is well known and published.



Trustworthiness

- **All data received are archived**
 - Includes water scenes and 100% cloud covered scenes as future uses may want access to these scenes
- **ACRES is in the early stages of archiving processed imagery for other groups within GA. These data are stored so that the original is precisely replicated when retrieved.**

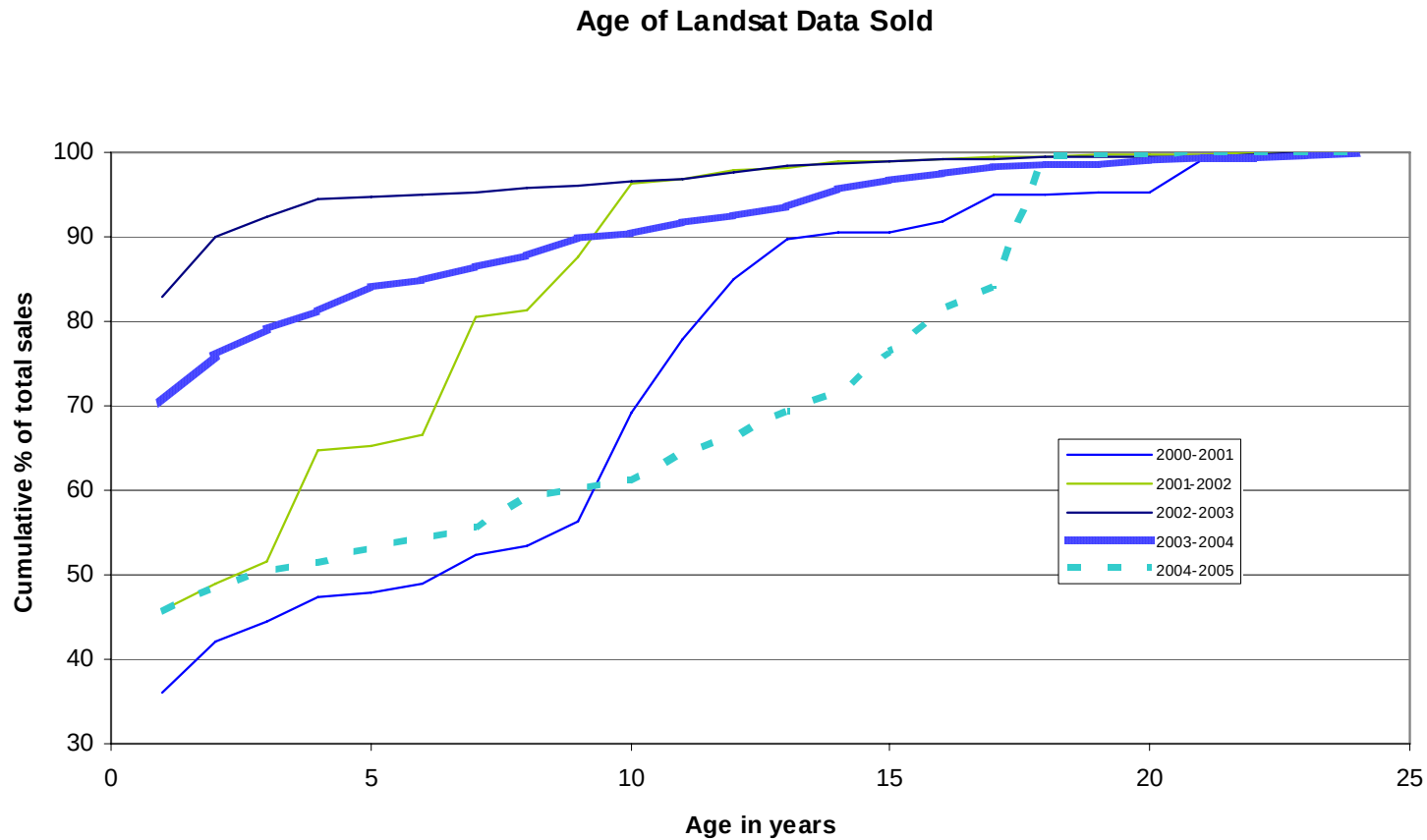


Value of the archive

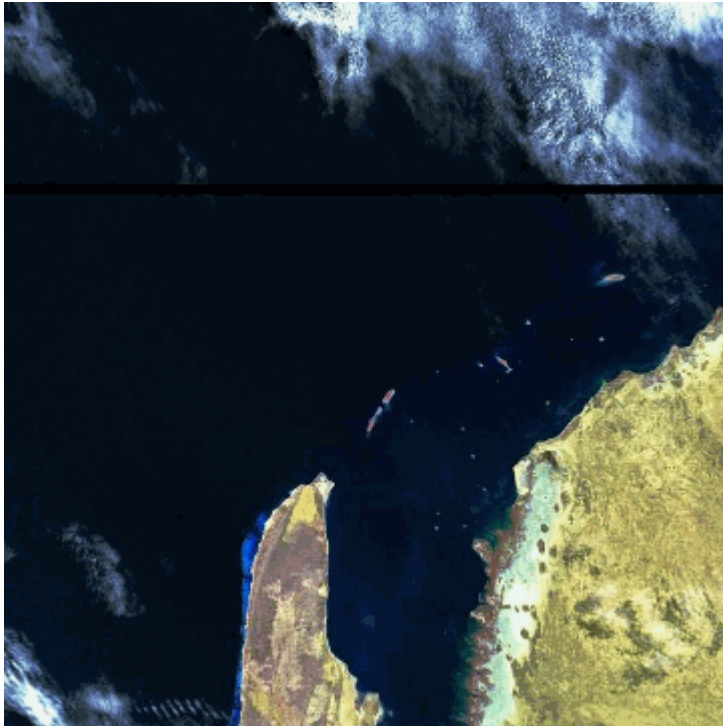
- The value of the archive is incalculable but very significant
- It represents the only continent-wide record of land cover spanning the last 25+ years
- Much of our archive is unique
- We cannot predict to what uses these data will be put in the future, and as an irreplaceable record it is incumbent upon us to maintain them indefinitely
- The archive is of increasing value for climate change research
- A significant proportion of data sales are of “old” data demonstrating the value of the archive



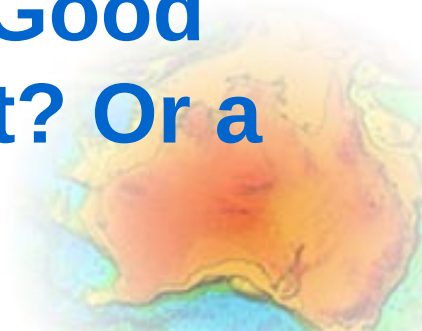
Value of the archive

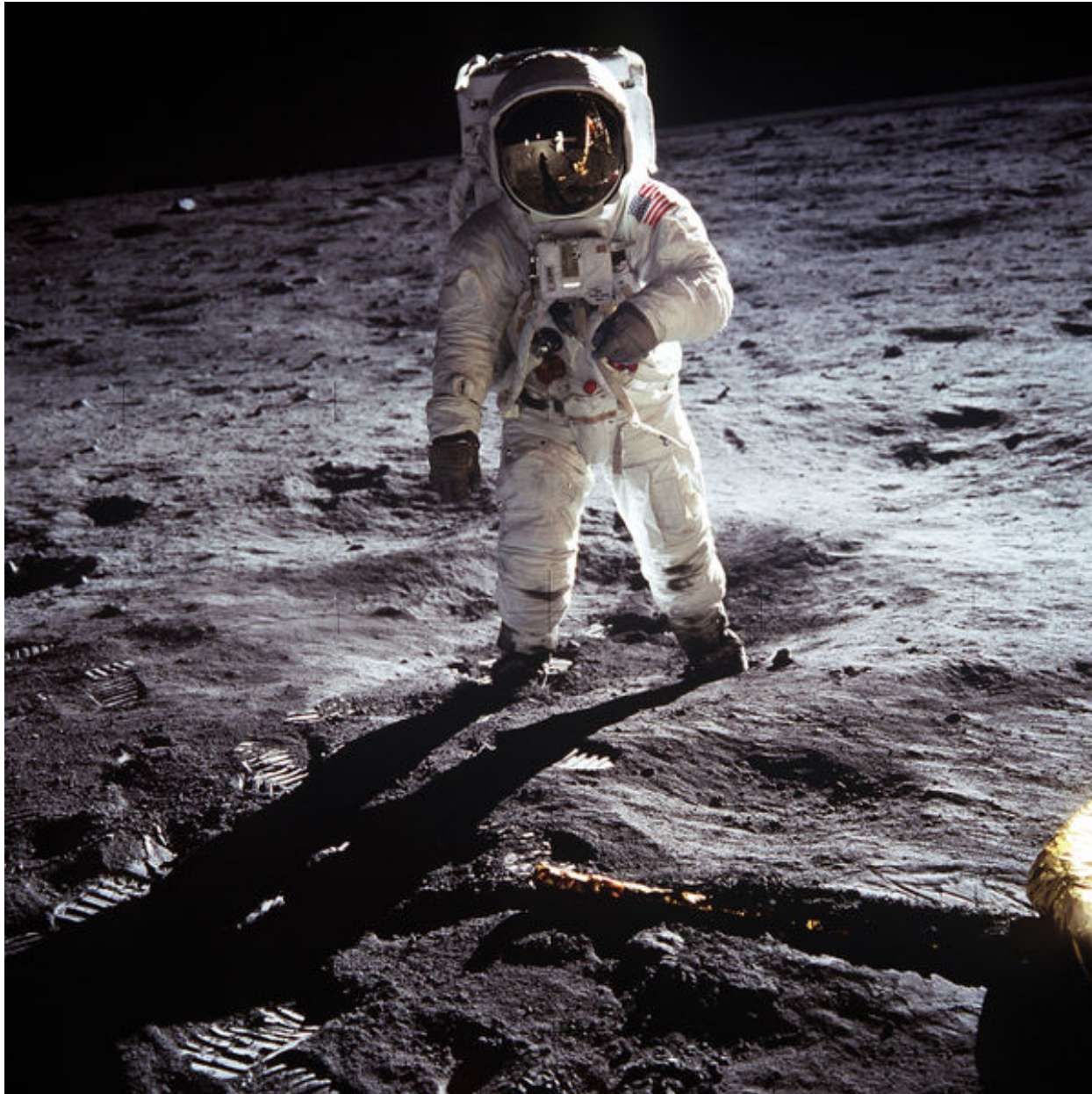


Data management



- Image is of North West Cape, WA
- Capture Date: 1979-09-14 01:34:06 UTC
- Data are still viable and accessible!
- Good luck? Good management? Or a bit of both?





NASA loses original tape of moon landing

August 15, 2006 - 8:20 am

The US government has misplaced the original recording of the first moon landing, including astronaut Neil Armstrong's famous "one small step for man, one giant leap for mankind," a NASA spokesman says.

Source – Sydney Morning Herald



Data management

- **ACRES decided long ago that the archive was a major asset and it was treated appropriately with sound management practices adopted including the maintenance of an off-site backup copy of all data and periodic updates of the storage medium**
- **Active management of the archive is necessary to maintain it into the future**



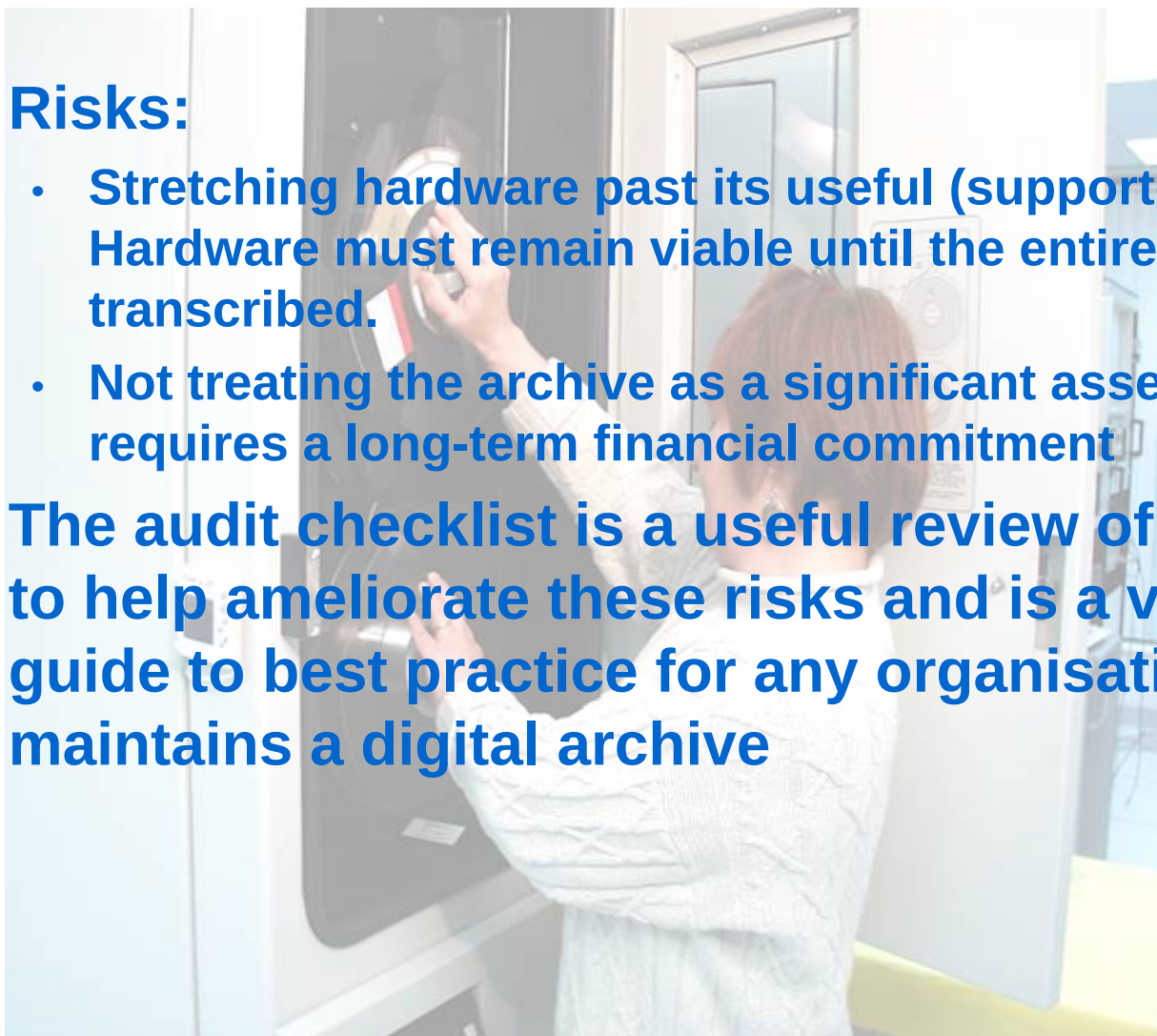
The next 25 years?

- **Challenges for the future include:**
 - **Rapidly growing archive volumes, particularly as new missions frequently increase data rates**
 - **Users' desires for on-line access**
 - **Metadata standards**
 - **How to manage distributed archives**
 - **Shorter manufacturer support periods for hardware requiring more frequent refreshes with concomitant cost and time implications**



The next 25 years?

- **Risks:**
 - Stretching hardware past its useful (supported?) life. Hardware must remain viable until the entire archive is transcribed.
 - Not treating the archive as a significant asset which requires a long-term financial commitment
- The audit checklist is a useful review of practices to help ameliorate these risks and is a valuable guide to best practice for any organisation that maintains a digital archive



Audit checklist

- **A. Organisation**
 - These questions help identify the executive management's commitment to maintenance of the archive
 - Although A1.1 and A1.2 are not directly relevant unless storing data on behalf of others, they are worthy of careful consideration for all archives
 - Good change control management (A3.4) is critical as is management of copyright (A5.3)



Audit checklist

- **B. Repository functions, processes & procedures**
 - A clear understanding of exactly what will be preserved (B1.1) is a foundation for design of systems and processes
 - Sufficient knowledge of the data structure and maintenance of systems required to turn bits into information is critical (B3)
 - Data are only useful if it can be found by users! Good metadata are necessary (B4)



Audit checklist

- **C. The designated community & the usability of information**
 - The concept of “designated community” is a useful concept to ensure users are appropriately served (C1.1)



Audit checklist

- **D. Technologies & technical infrastructure**
 - Key section for ensuring data can be retrieved in years to come
 - Rather than referring to media life in D1.7, it would be more reasonable to consider hardware support as the trigger for migration



Lessons learned

- **Hardware support for archive media is critical**
 - Hardware problems can compromise access to the archive well before the expected lifetime of the media
 - Platform dependencies in data structures must be avoided
- **Comprehensive records of where each particular record is held are necessary; these should be systematic and automatic**
- **Extensive metadata are required**
- **Easily searchable catalogues improve accessibility to data**
- **Archives need to be treated as a valuable asset with a planned maintenance strategy**





Thank you!