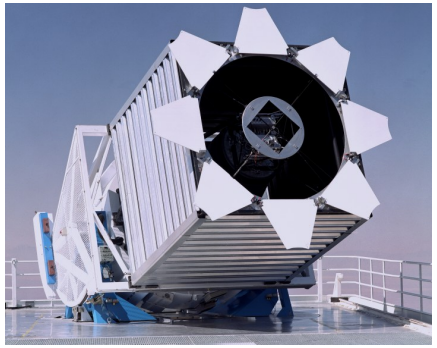
Lessons from MaNGA and SDSS-IV



Anne-Marie Weijmans
University of St Andrews

University of St Andrews

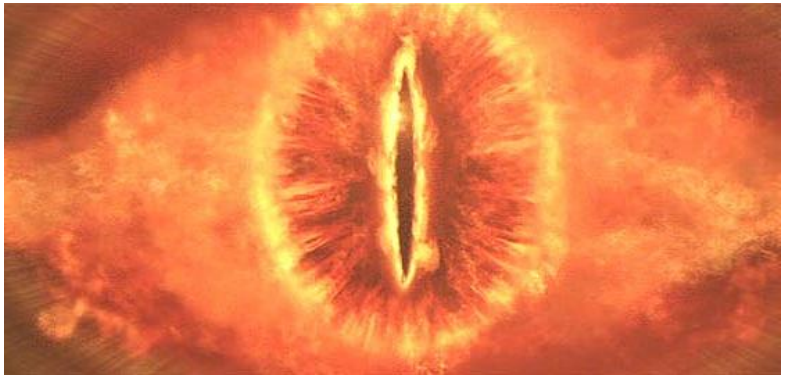
My astronomy career in surveys

Surveys Past:

- SAURON survey (undergraduate / PhD student)
- Atlas3D survey (PhD student / postdoc)

Surveys Present:

- Atlas3D/Mitchell survey (postdoc)
- SDSS-IV (postdoc / staff)
  - MaNGA lead observer
  - SDSS-IV data release coordinator
  - co-lead of SDSS-IV data team

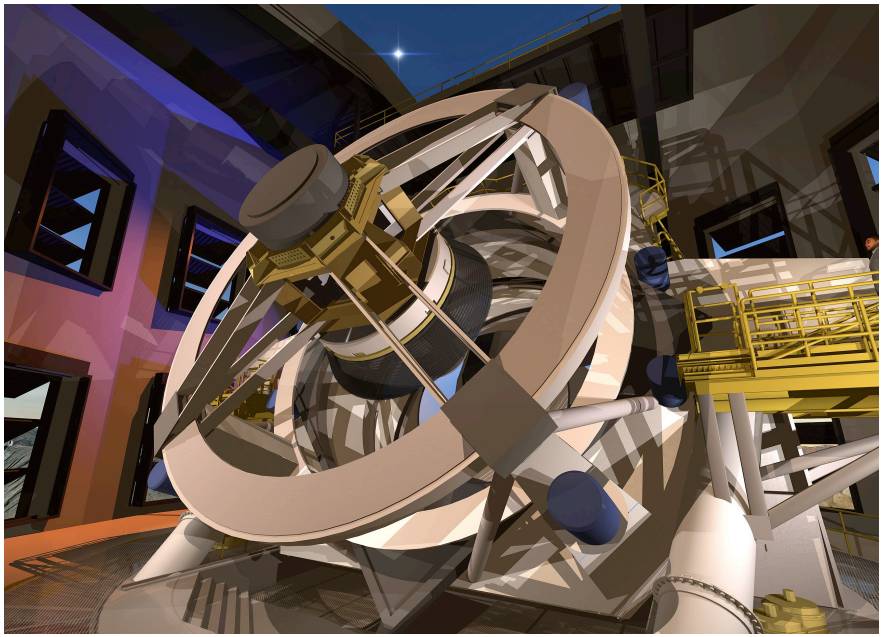


Surveys Future:

- WEAVE
- SDSS-V

Why surveys?

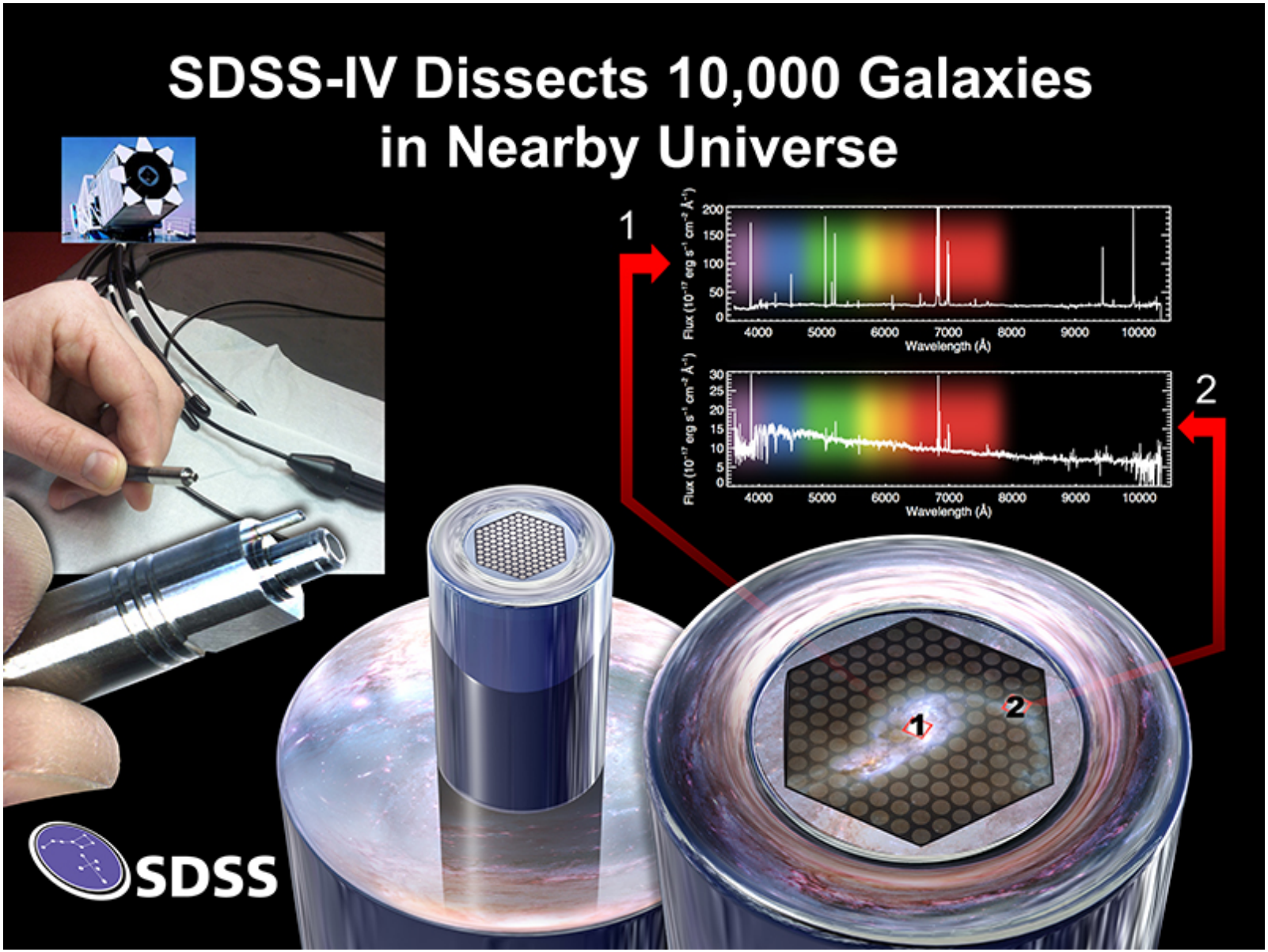
- Big questions need big resources



Credit: LSST

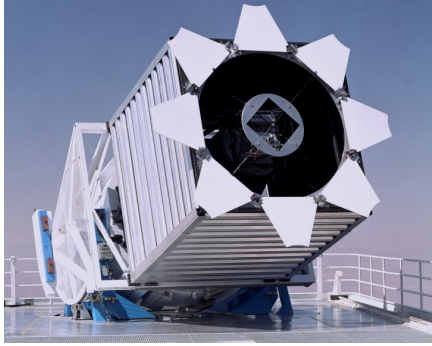
Lessons from MaNGA and SDSS-IV

- Introducing MaNGA
  - survey design and goals
  - science highlights
- How does a large collaboration do science?
  - examples from MaNGA and SDSS-IV
- What makes a successful collaboration?
  - and what makes them fail?
- Should *you* join a large collaboration?
  - some questions to ask yourself


SDSS-IV Dissects 10,000 Galaxies
in Nearby Universe
1
2
Flux (10⁻¹⁷ erg s⁻¹ cm⁻² Å⁻¹)
Wavelength (Å)
SDSS

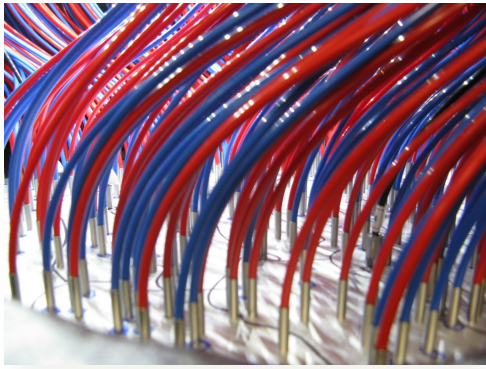
Bundy et al. 2015

Key Questions



- Lives of galaxies
  - How does gas accretion drive the growth of disks?
  - What are the relative roles of stellar accretion, major mergers and instabilities for forming bulges?
- Death of galaxies
  - What quenches star formation?
  - How is star formation affected by groups and clusters?
- Birth of galaxies
  - What was the initial distribution of angular momentum?
  - How do baryons and stars trace and influence the dark halo?
  - Is galaxy growth the same at low and high z?

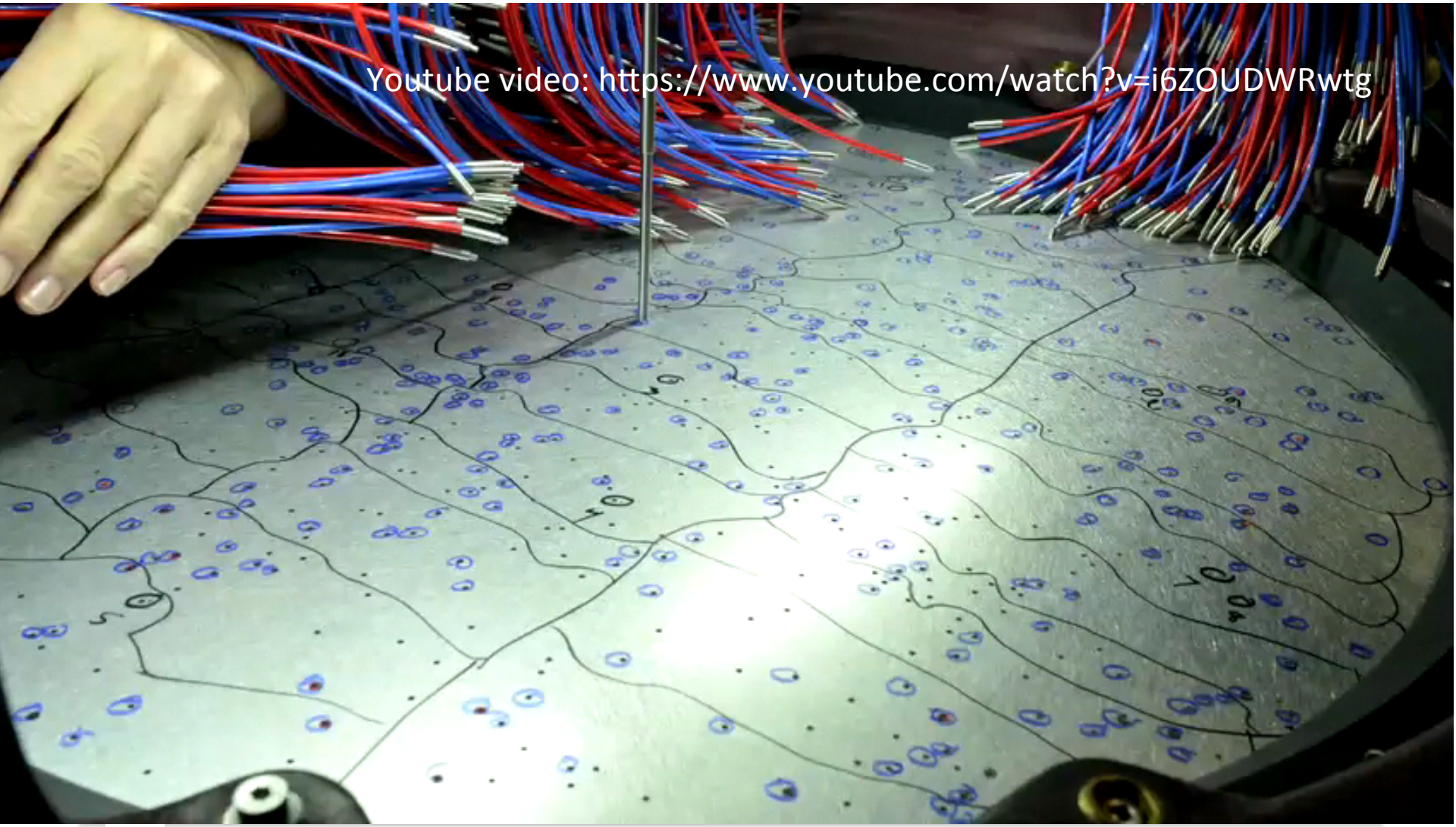
MaNGA Concept



- Bundle single fibers together in IFUs
  - use BOSS spectrographs
  - create 17 IFUs, from 19 to 127 fibers per bundle
- Plug bundles in plates, similar to single fibers
  - integrate sky fibers with IFUs
  - 12 mini-IFUs for standard stars (7 fibers each)
- Observe 10,000 galaxies, 3hr dithered exposures
  - spatial resolution: 2" fibers or 1 – 4 kpc
  - spectral resolution: 50 – 80 km/s (R = 2500)
  - spectral coverage: 3600 – 10,000 A
  - S/N: ~ 30 in central fiber, ~4-8 at 1.5 R_e

Bundy et al. 2015
Weijmans et al. 2016

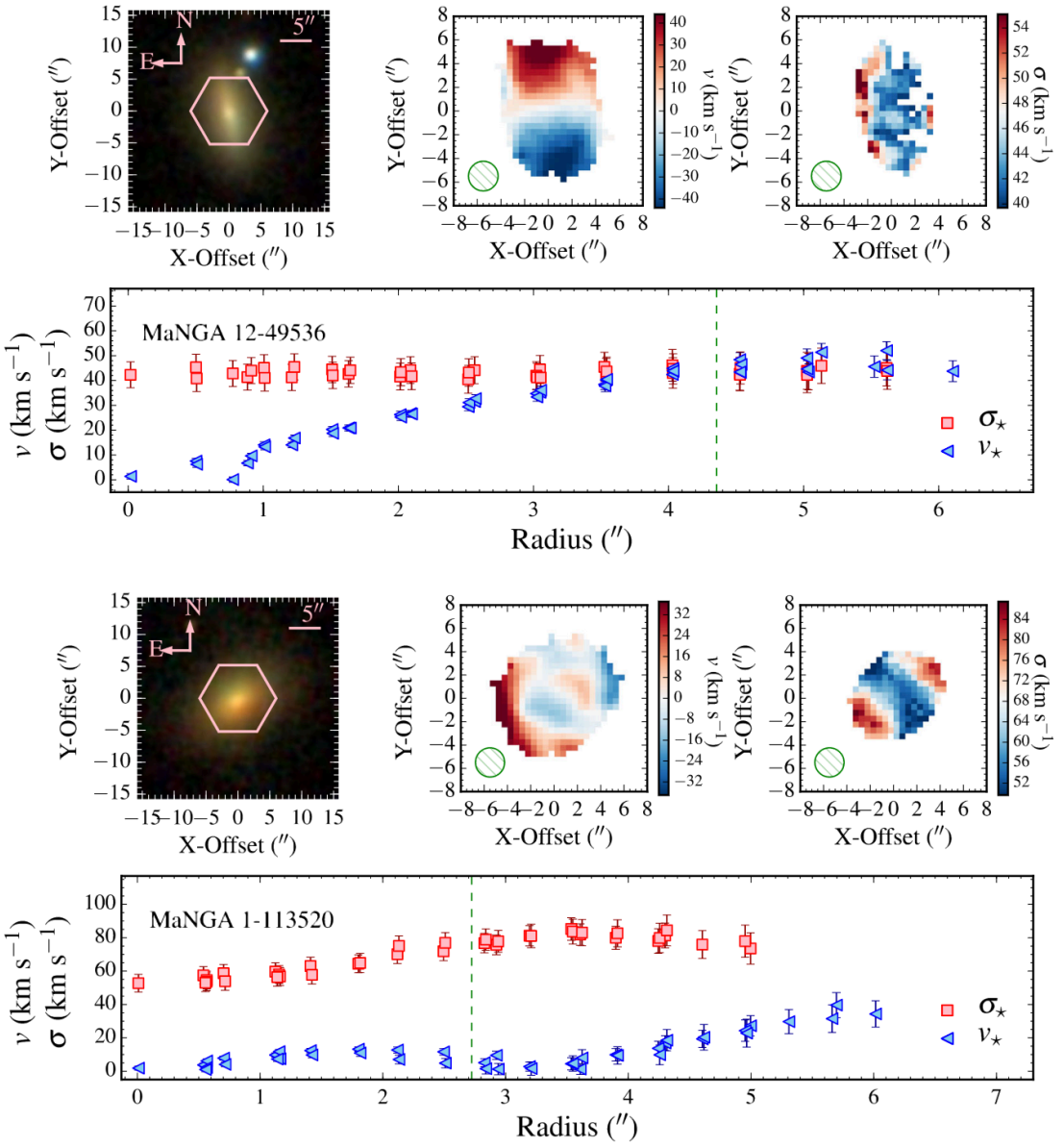
Plugging a plate



Sloan Observing

Youtube video: https://www.youtube.com/watch?v=AHsS57NMQjE

MaNGA: dwarf galaxies

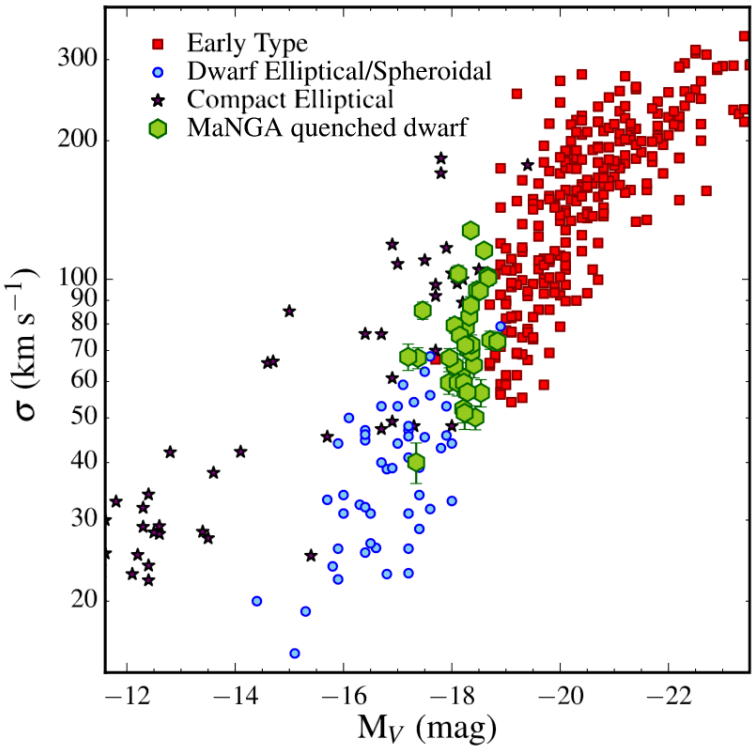


Penny et al. 2016

MaNGA: dwarf galaxies

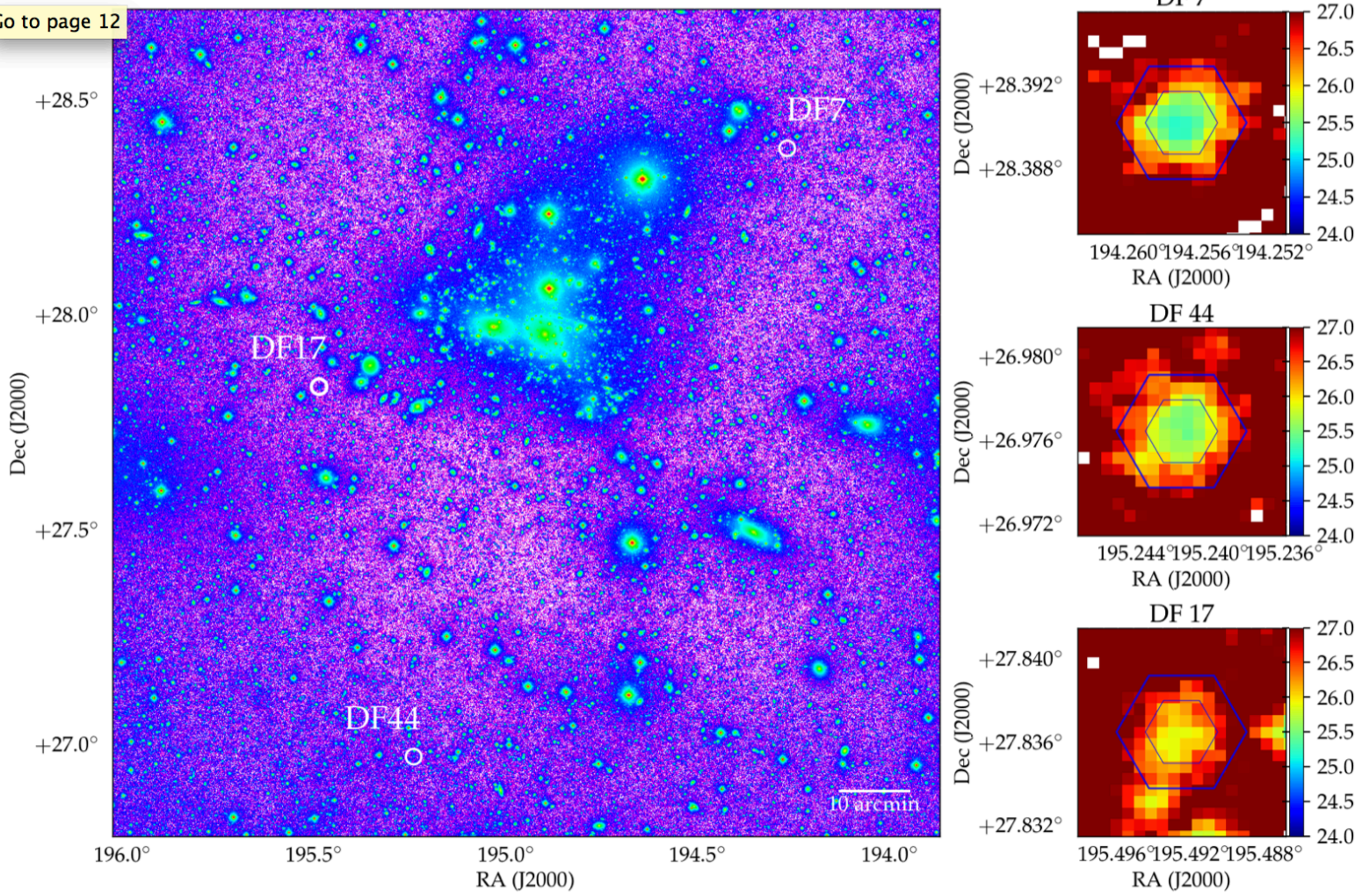
- Only 6/39 (~15%) have rotation < 15 km/s
  - these are pressure supported at all radii
- Two have kinematically distinct cores
- Sample follows σ-luminosity scaling relations of early-types

Penny et al. 2016

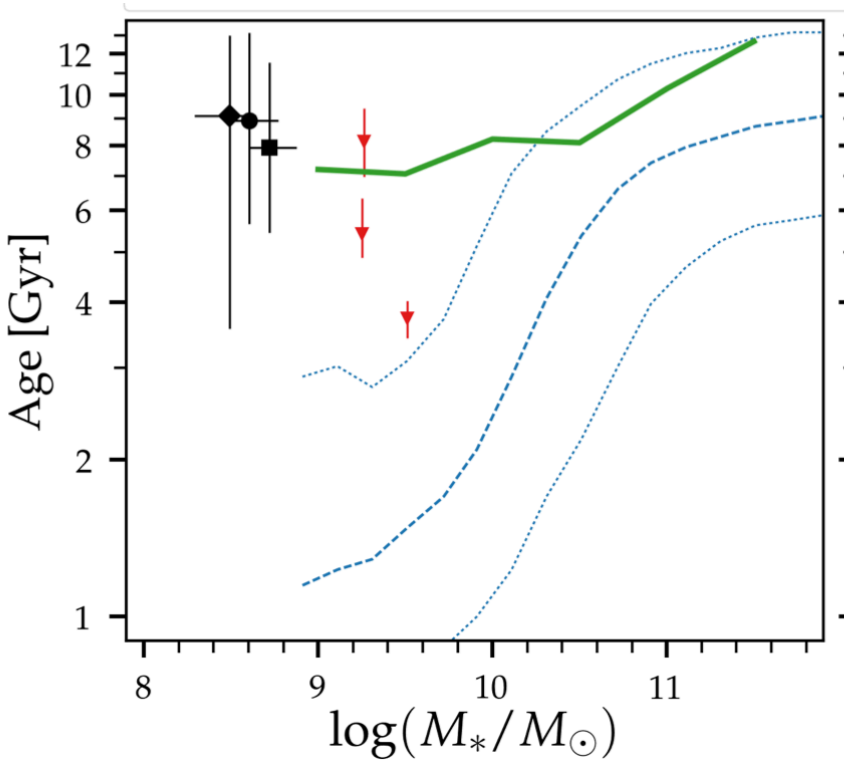


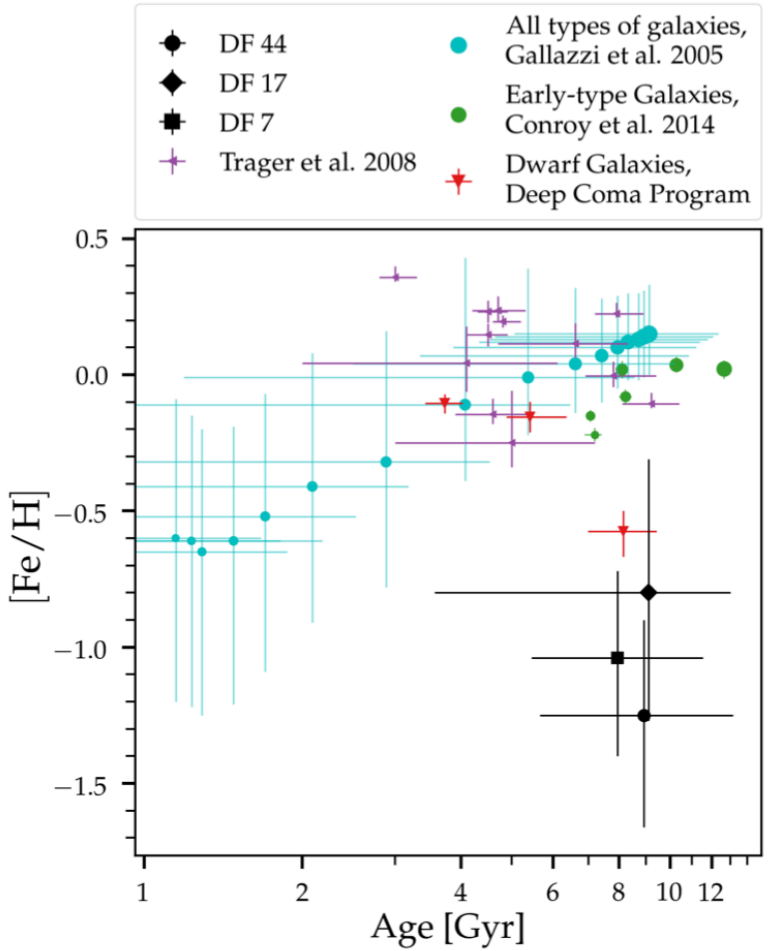
MaNGA: Coma Cluster

Gu et al. 2018



Gu et al. 2018



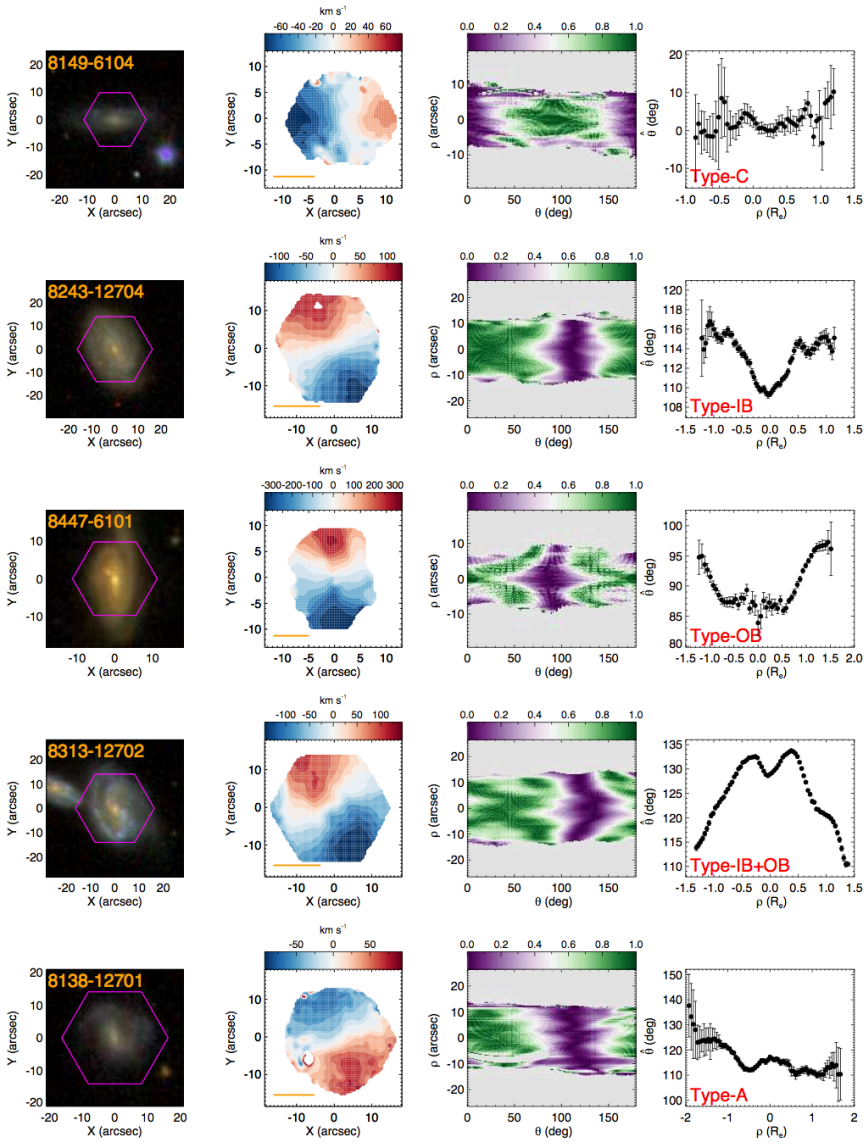


- UDGs are old and metal-poor
  - slightly older than ETGs → not recently quenched
  - brief SF period at high redshift

MaNGA: position angles of velocity fields

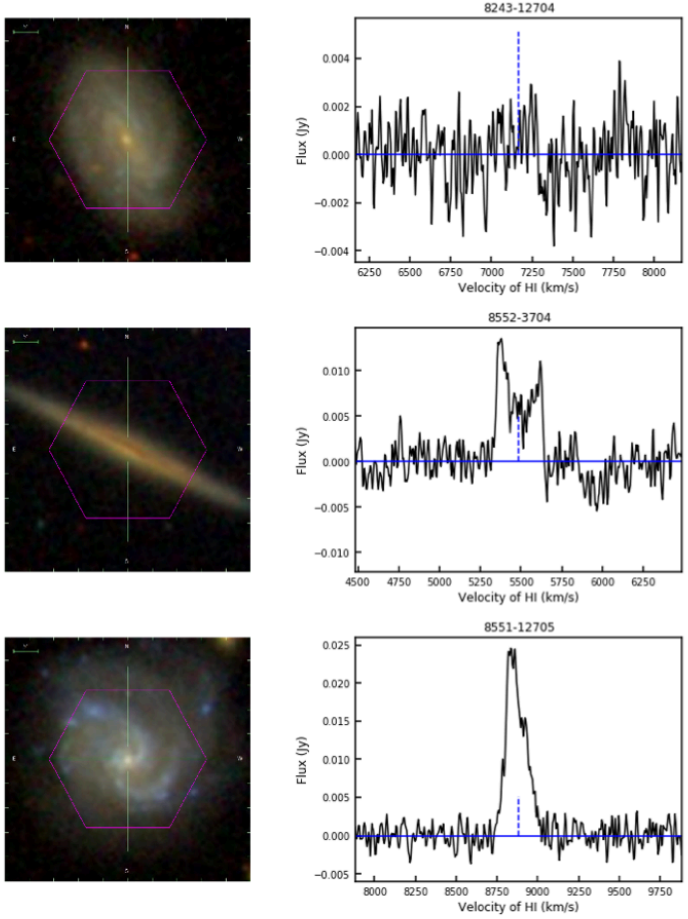
- Uses radon method to characterize position angle profiles of velocity fields

Stark et al. 2018



Masters, Stark, et al. in prep

MaNGA and HI

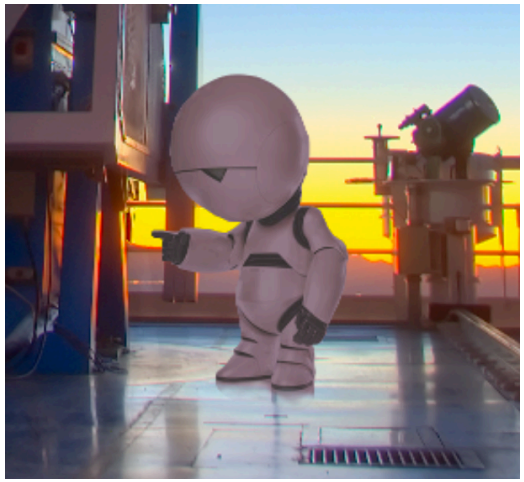


- 21 cm follow-up with Greenbank Telescope
- 331 galaxies observed so far
- additional data from ALFALFA
- value added catalog in preparation (available this December in DR15)

Next SDSS-IV Data Release: DR15

- More MaNGA data cubes!
- New: MaNGA maps!
- New: MaStar stellar library stars!
- New: Marvin!

December 2018

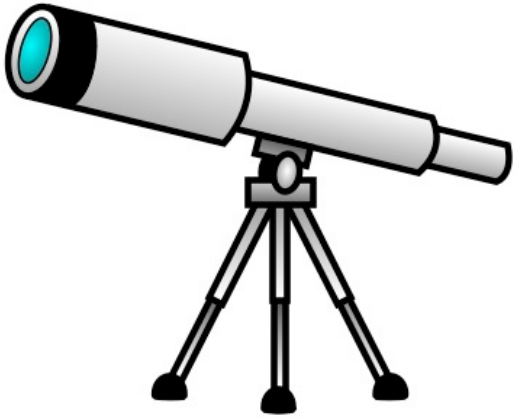


www.sdss.org

How does a large collaboration science?

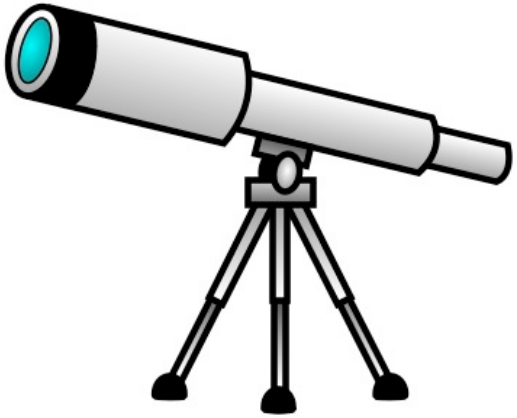
How does a large collaboration science?

- Infrastructure

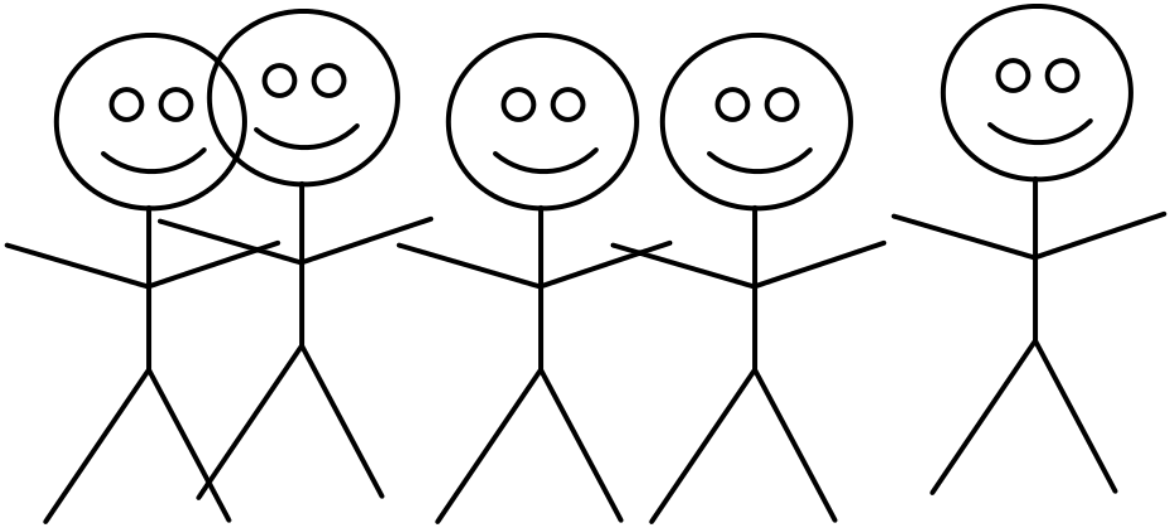


How does a large collaboration science?

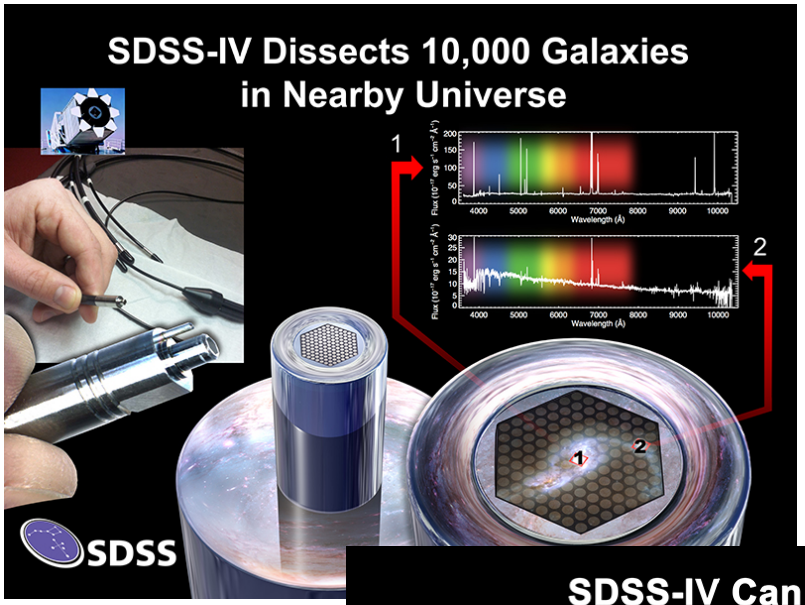
- Infrastructure



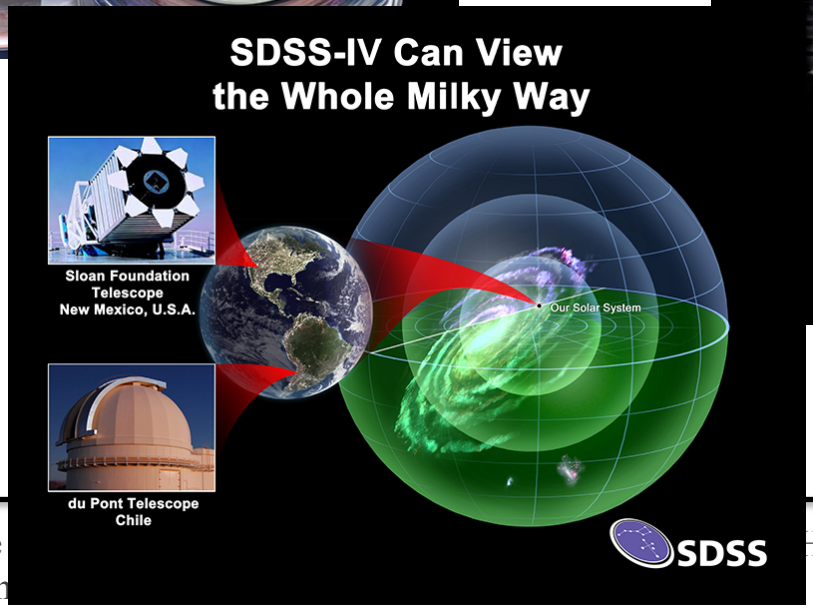
- People



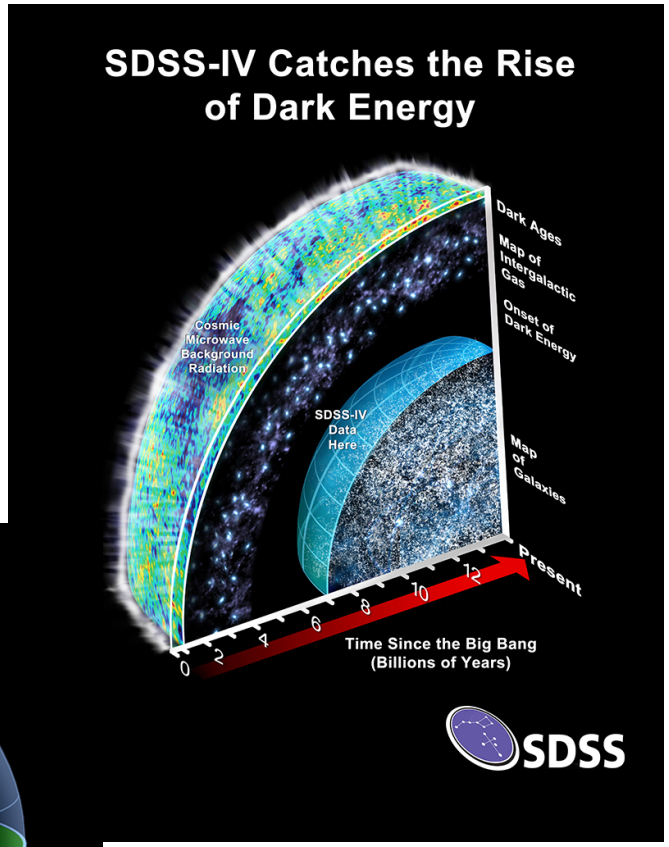
Sloan Digital Sky Surveys IV



MaNGA



APOGEE-2



eBOSS

What does SDSS deliver?

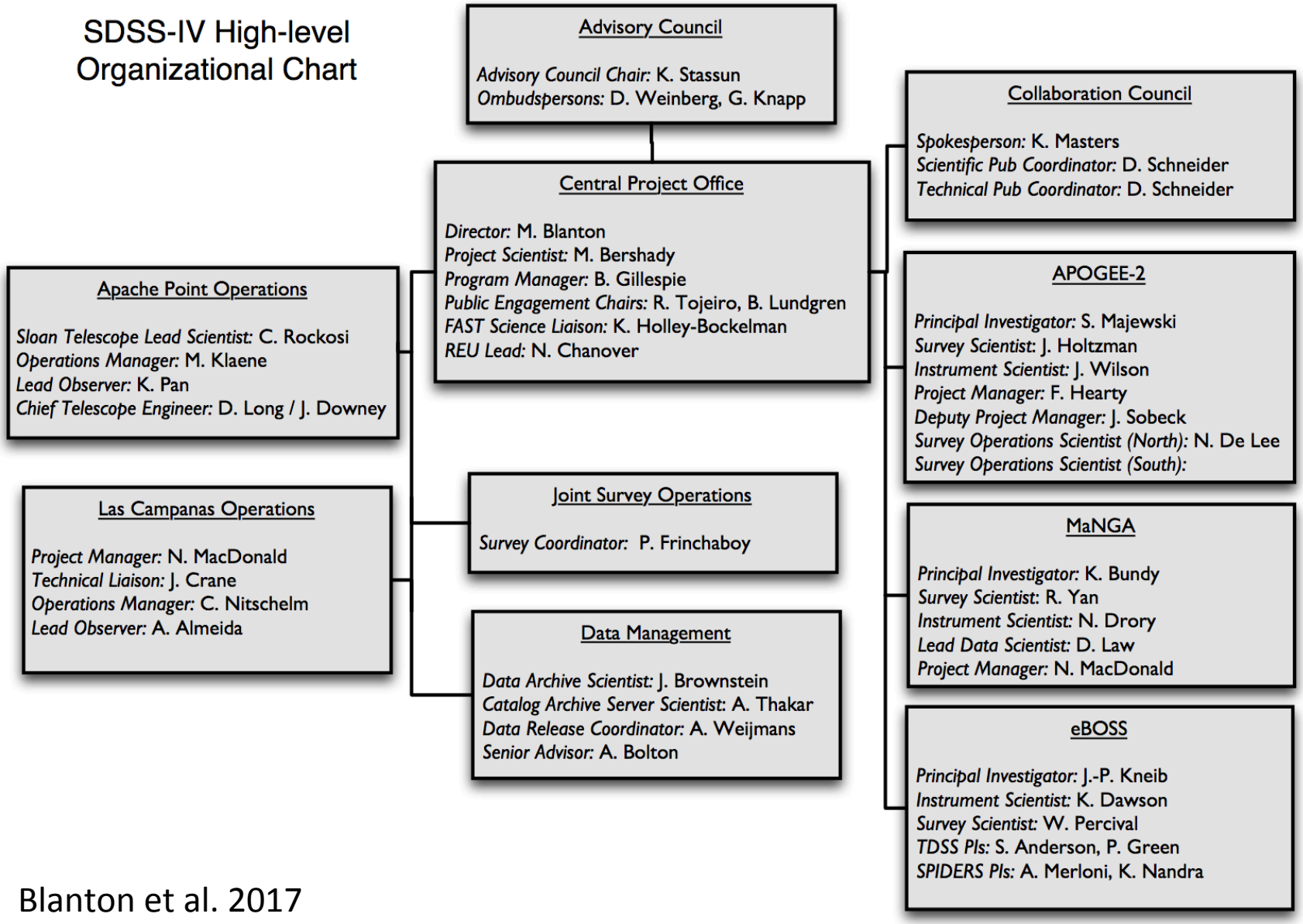
Weijmans et al. 2016, arXiv: 1612.05668

- Data publicly available
  - data release 14 (July 2017) contains 325 TB
- Education and public engagement activities
  - Voyages, plates program (www.sdss.org/education)
  - citizen science (galaxy zoo)
  - artist in residence (based in St Andrews)
- Very high impact survey
  - more than 7000 papers based on SDSS data
  - 30% of US community uses SDSS data
  - trusted resource for teachers and general public

SDSS infrastructure

- Survey Design
  - science questions, science requirements
  - targets, observing strategy
- Observations at Telescope
  - maintenance, operation software updates
- Data reductions and analysis
  - data transferred from mountain to data centre
  - data pipelines run on dedicated SDSS servers
  - data products vetted and documented by data teams
  - distributed to science teams
- Science happens
  - papers
  - value added catalogs

SDSS-IV High-level Organizational Chart



Blanton et al. 2017

SDSS organisation

- Surveys are organised in working groups
  - regular telecons
  - team meetings
- Projects, papers and value added catalogs are announced on mailing lists
  - wiki pages to share progress on projects etc.
- 'You can work on anything that you want, as long as you tell people'
  - external collaborator requests are possible
- Code of Conduct

How does a survey define success?

How does a survey define success?

- Papers written by survey team members?

How does a survey define success?

- Papers written by survey team members?
  - do papers written with public data count too?

How does a survey define success?

- Papers written by survey team members?
  - do papers written with public data count too?
- Terabytes of data generated?
- Value added catalogs based on survey data?

How does a survey define success?

- Papers written by survey team members?
  - do papers written with public data count too?
- Terabytes of data generated?
- Value added catalogs based on survey data?
- Survey data used by teachers?
- Survey data used for public engagement?

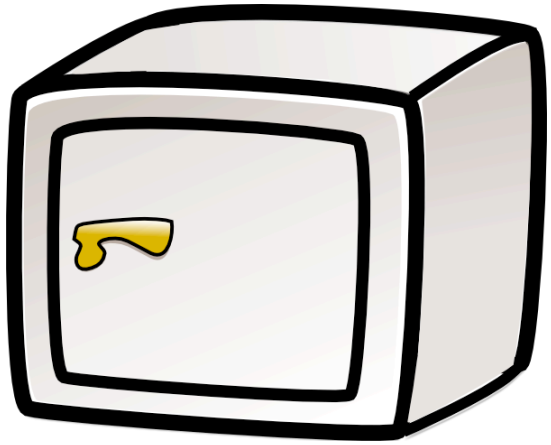
How does a survey define success?

- Papers written by survey team members?
  - do papers written with public data count too?
- Terabytes of data generated?
- Value added catalogs based on survey data?
- Survey data used by teachers?
- Survey data used for public engagement?
- Press coverage?

How does a survey define success?

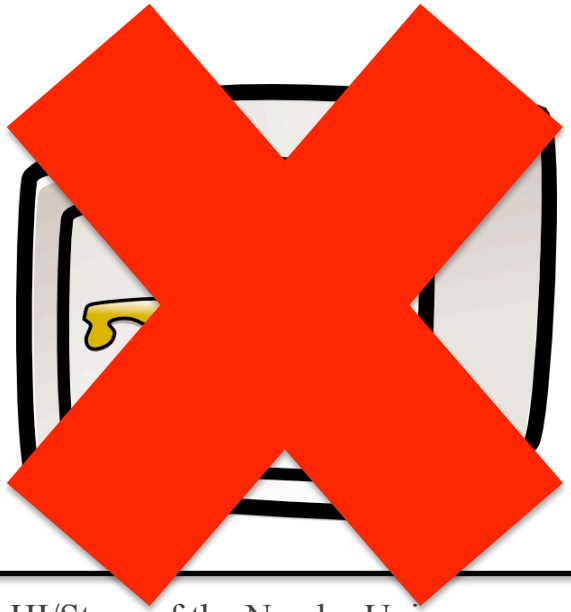
- Papers written by survey team members?
  - do papers written with public data count too?
- Terabytes of data generated?
- Value added catalogs based on survey data?
- Survey data used by teachers?
- Survey data used for public engagement?
- Press coverage?
- Career advancement for team members?

Make your data accessible!



Make your data accessible!

- My pet peeve: available ≠ accessible
- Each survey needs:
  - where is which data product?
  - data models
  - list of caveats, bugs
  - tutorials
- Not accessible just for public, but also for your own team!



Side note: don't underestimate public data releases…

- DR13 happened July 2016
  - we started prepping in October 2015

Weijmans et al. 2016

- Long process involving whole SDSS team
  - identify core data products for release
  - quality control of data products to be included in release
  - call for value-added catalogs
  - identify software to be posted in SVN depository
  - identify documentation needs
  - prepare and load data into SAS → science archive
  - prepare and load data into CAS → catalog archive
  - construct DR13 website, write DR13 paper
  - mirror archive servers, back-up on HPSS tapes
  - test servers, install metric evaluations
  - organise social media
  - all the things you forgot that pop up in the last month
  - actual data release

How does a survey define success?

- Papers written by survey team members?
  - do papers written with public data count too?
- Terabytes of data generated?
- Value added catalogs based on survey data?
- Survey data used by teachers?
- Survey data used for public engagement?
- Press coverage?
- Career advancement for team members?

Be nice!

Be nice!

- Provide a friendly, safe environment
  - be welcoming to new team members
  - have a Code of Conduct and ombudspersons
- Be inclusive
  - aim for a diverse survey team AND leadership
- Provide opportunities and encouragement for career growth
  - mentoring, early-career networks, deputy positions
- Thank people for their work!

Examples in SDSS

- Code of conduct
  - https://www.sdss.org/collaboration/code-of-conduct/
- Committee on Inclusiveness (COINS)
  - https://www.sdss.org/collaboration/coins/
- FAST and REU programs
  - https://www.sdss.org/education/

What can make a survey fail?

- Lack of communication
- Lack of person power
- Rules
- People

Lack of communication

- Keep all lines of communication open
  - mailing lists, wiki
  - telecons, team meetings
  - accessible data and documentation
- Be transparent
- Be inclusive
  - advertise opportunities
  - beware the time zone bias!
- Be welcoming
  - lower barriers for new team members

Lack of person power

- Is there enough person power to keep the survey going? → be realistic
  - do people have the time and resources they need?
  - what are the priorities?
  - don't overpromise!
- Is all knowledge shared within the team?
  - if someone leaves the team, will the survey collapse? (single-point failures)
- Do survey contributions get acknowledged?
  - builder or architect status?

Too strict rules on collaborations

- Can team members still work with their friends who are not in the team?
  - can team members submit external collaborator requests?
  - if yes, what are the rules?
- Sharing data with another survey?
  - survey-wide agreement?
- You don't want to limit science!

The rogue scientist

- Someone thinks the rules don't apply to them…
  - what action should the survey take?
  - what action can the survey take?
  - what is the legal framework?
- Abusive behavior should never be tolerated!

Should *you* join a large survey?

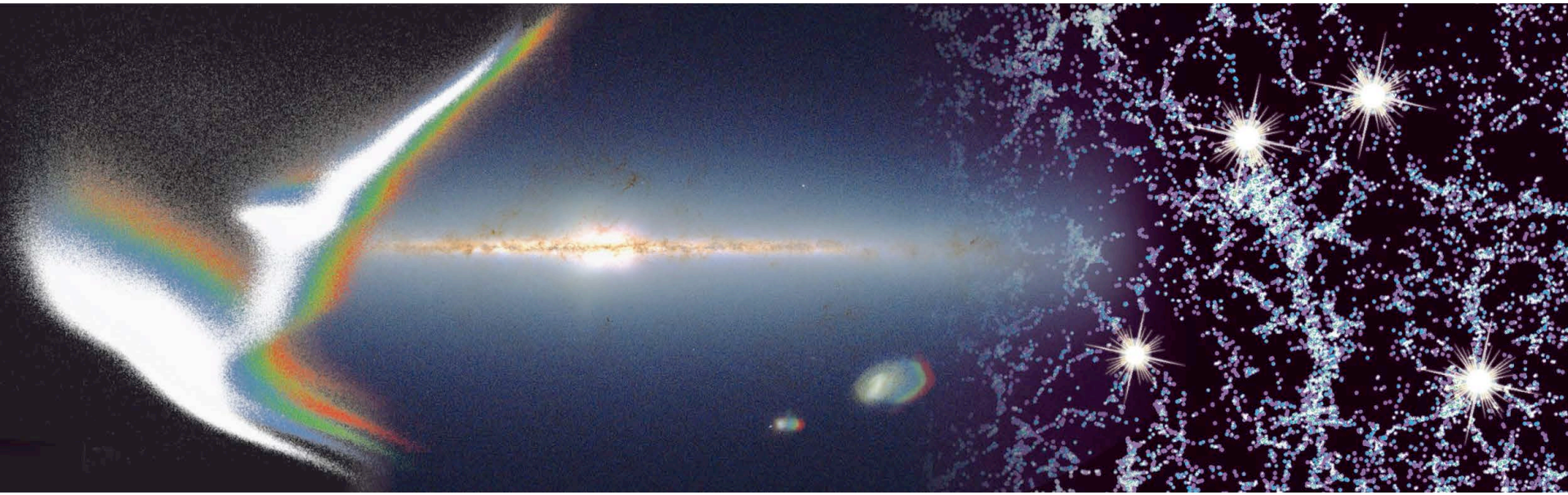
- Does the survey meet your goals?
  - or should you start your own one?
- Do you want to work in this survey?
  - too big, too small, too formal, too informal?
- Find out how to become a team member
  - are you already a member through your research council, institute or supervisor?
  - membership fee, work-load?
  - talk to people!

Do not get lost!

- Know what you want to get out of the survey
  - science, transferable skills, leadership?
- Learn how the survey is structured
  - attend telecons, meetings, talk to people
- Find out about career support
  - mentoring, early career networks?
- Find your niche
  - do your thing!

Sloan Digital Sky Surveys V
Pioneering Panoptic Spectroscopy

Multi-object Spectroscopy: all-sky, optical and near-IR, multi-epoch
Integral Field Spectroscopy: 3000 deg^2, optical, ultra wide-field



White paper:
Kollmeier et al. 2017, arXiv:1711.03234

Webpage:
http://www.sdss.org/future/

My conclusions

- Large surveys are fruitful for science
  - we can all benefit directly or indirectly
- Make survey data publicly accessible
  - available ≠ accessible
  - good documentation is key
- Engage with the wider community
  - impact beyond academia with public engagement
- Be nice, inclusive, and value your colleagues