Spectroscopic Surveys: Future

Mark Cropper
Mullard Space Science Laboratory,
University College London
Contents

• Future large surveys
• Features in common
• Current VO capability
• What we might like to do in the future
• Summary

thanks for inputs from

Will O’Mullane, Nic Walton, Pedro Osuna, Tamas Budavari & Ivo Busko
Future Surveys

• Spectroscopic surveys will grow in size, with increasing capability:
  – satellite all-sky surveys
  – multi-object spectrographs with >1000 fibres
  – large integral field spectrographs
  – major new sub-mm and radio facilities (ALMA, SKA)

• Criteria adopted here is that the survey should be >100000 sources

• Large surveys fall into three broad categorisations:
  – dark energy surveys (baryon acoustic oscillation or BAO surveys)
  – galaxy formation
  – structure of our Galaxy

… take these in turn
Dark Energy Surveys: WiggleZ

- 400000 star-forming galaxies
- 400 spectra/exposure
- limiting magnitude $20.5 < r < 22.5$
- facility: AAT + AAOmega
- spectral range 370-900 nm
- resolving power $R \sim 1-2000$
- starting $\sim 2006$ ending 2009
- approved, under way

Dark Energy Surveys: FastSound

- 600000 star-forming galaxies (Hα)
- 400 spectra/exposure
- limiting magnitude J ~ 22
- facility: Subaru + FMOS
- spectral range 0.9 - 1.8 μm
- resolving power R~500
- starting ~2008
- proposed, not yet approved

http://www-utap.phys.s.u-tokyo.ac.jp/~utap/meetings/workshop/hsc2006/English/totani.ppt
Dark Energy/Large Scale Structure Surveys: SDSS post-2008

- 1 million galaxies and Ly-α absorbers
- 640 per exposure
- limiting magnitude $g \sim 20$
- facility: Apache Point 2.5m telescope
- wavelength range 380-920 nm
- resolving power $R \sim 2000$
- starting 2008 completing 2012
- proposed

Dark Energy Surveys: LAMOST

- 10 million Luminous Red Galaxies
- 4000 spectra/exposure
- limiting magnitude V~20.5
- facility: LAMOST + LRS
- spectral range 370-900 nm
- resolving power R~2000
- starting ~2008
- approved

http://ej.iop.org/links/r_nx_nc4r/atq-vl3V2xGMKyGYav5vpA/chjaa_6_3_001.pdf
http://www.lamost.org/en/
Dark Energy Surveys: HEXDEX

- 1 million Ly-\(\alpha\) emitting galaxies
- 36000 spatial elements/exposure
- limiting magnitude AB<23.5
- facility: Hobby-Eberly Telescope + virus
- spectral range 340-570 nm
- resolving power R~850
- starting ~2010
- approved

http://www.as.utexas.edu/hetdex/Hill_MitchellSymposium_mod.pdf
http://www.as.utexas.edu/hetdex/
Dark Energy Surveys: Subaru/WFMOS

- 3 million galaxies $z = 0.5$ to $3.3$
- 4-5000 targets per exposure
- limiting mag $R_{\text{lim}}$ AB = 22.7
  (AB = 24.5 for 1/3 of objects)
- facility: Subaru+WFMOS
- spectral range 400nm within 390-10000nm
- resolving power 13500
- starting 2012?
- approved (on hold)

http://www.noao.edu/meetings/subaru/Session-1/Colless.pdf
Dark Energy Surveys: ADEPT

- 100 million galaxies \(1 \leq z \leq 2\)
- slitless spectroscopy
- limiting magnitude AB<?
- facility: ADEPT satellite
- spectral range \(~ 1.3 - 2\mu m\)
- resolving power R~?
- starting after \(~2015\)?
- 1 of 3 contenders selected for NASA JDEM

http://www7.nationalacademies.org/ssb/mtg_2_ADEPT.pdf
Dark Energy Surveys: SKA

- $10^8 - 10^9$ galaxies to $z \sim 3$ (HI)
- large numbers of beams
- limiting sensitivity: 100 x VLA
- facility: SKA
- spectral range 100 MHz - 25 GHz
- 10000 spectral channels
- starting ~2019
- under study

http://www.skatelescope.org
Galactic Structure: RAVE

- 1 million stars
- 150 spatial elements/exposure
- V~13.5 (some fainter)
- facility: UKSchmidt+6dF
- wavelength range 850-875 nm
- resolving power R~10000
- started 2003 (pilot); 2006-2010 (main)
- approved, under way

http://www.rave-survey.aip.de/rave/pages/project/ProjectDescription_1.jsp
Galactic Structure: SDSS SEGUE

- 250,000 thick disk/halo stars
- 640 per exposure
- limiting magnitude g ~ 20
- facility: Apache Point 2.5m telescope
- wavelength range 380-920 nm
- resolving power R~2000
- completing 2008
- approved, under way

http://www.sdss.org/segue/aboutsegue.html
Galactic Structure: ARGOS

- 110000 stars in bulge
- 400 spectra/exposure
- limiting magnitude $I_C \sim 17$
- facility: AAT + AAOmega
- spectral range 525-560 and 846-886 nm
- resolving power $R \sim 10000$
- starting ~2007 ending 2009
- proposed

Galactic Structure: SDSS post-2008

- 2 million stars
- 640 per exposure
- limiting magnitude $g \sim 20$
- facility: Apache Point 2.5m telescope
- wavelength range 380-920 nm
- resolving power $R \sim 2000$
- starting 2008 completing 2014
- proposed

Galactic Structure: LAMOST

- 10 million stars
- 4000 spectra/exposure
- limiting magnitude $V \sim 20.5$
- facility: LAMOST + LRS
- spectral range 370-900 nm
- resolving power $R \sim 2000$
- starting ~2008
- approved

http://ej.iop.org/links/rUR2c8GjO/ZPqta1zV2xG-XFiTav5vpA/chjaa_6_2_003.pdf
http://www.lamost.org/en/
Galactic Structure: Gaia

- 100 million stars
- ~40 exposures per star
- V~17.5 (some fainter)
- facility: satellite
- wavelength range 847-874 nm
- resolving power R~11000
- starting 2012 completing 2017
- approved

http://www.rssd.esa.int/GAIA/
Galaxy Formation: KMOS

- ?? galaxies
- 200 x 24 spatially resolved spectra/exposure
- K \sim 21
- facility: VLT + KMOS
- wavelength range 0.8 - 2.5 \mu m
- resolving power R \sim 3500
- starting \sim 2011
- details unclear

http://www.eso.org/instruments/kmos/#Science
Galaxy Formation: MUSE

- ?? galaxies
- 90000 spatially resolved spectra/exposure
- limiting magnitude $I_C \sim 25.0$
- facility: VLT + MUSE
- wavelength range 465 - 930 nm
- resolving power $R \sim 3000$
- starting ~2011
- details unclear

http://www.eso.org/instruments/muse/
Galaxy Formation: Herschel

- 850 hr of SPIRE instrument time
- 650 hr of PACS instrument time
- ~230 mJy (continuum) (SPIRE)
- facility: Herschel + SPIRE + PACS
- wavelength range 60 - 670 µm
- resolving power R ~ 300-1200 (SPIRE)
- starting ~2008

http://sci.esa.int/science-e/www/object/index.cfm?fobjectid=34682
http://www.astro.cf.ac.uk/groups/instrumentation/projects/spire/Public_Index.htm
Other important spectroscopic surveys

- XMM-Newton/Chandra X-ray surveys continue to increase in coverage
- AKARI (Astro-F) infra-red survey
- ALMA (sub-mm)
- LOFAR (radio)

etc…
Features in Common: observational characteristics

• **Scale**: from $10^4$ to $10^8$ sources (some observed many times); largest are SKA, Gaia-RVS, ADEPT, LAMOST and Subaru+WFMOS (>few million)

• **Spatial connectedness**: some are all-sky, such as Gaia-RVS, others are spatially closely connected (VLT-MUSE; VLT-KMOS)

• **Spectral resolving power**: generally low-moderate ($R \sim 500-10000$); no large high resolution surveys

• **Wavelength range**: most of the largest spectral surveys are in the optical/radio, but significantly sized surveys also at other wavelengths (eg XMM-Newton/Chandra).

• **Signal-to-noise ratios**: most surveys will be exposed for reasonable S/N ratios, but the biggest (Gaia-RVS, ADEPT, SKA) will have data with a whole range of S/N ratios, the majority very low
Features in Common: observational characteristics

- 3 orders of magnitude increase in 10-15 years!
A little more on *Gaia*-RVS

- *Gaia* will produce a total of 4 billion spectra (100 million x 40 transits)
- most of the data will be very low S/N ratio
- both single transit and multiple-transit (*i.e.* combined) data will be available
- some derived products will be available:
  - \( V_{\text{rad}} \), \( V_{\text{rot}} \)
  - classifications,
  - indications on binarity/multiplicity; if so a measure of the period
    but probably not line strengths, metallicities, gravities *etc.*
- VO-enabled via team here at ESAC

- Significant need and many opportunities for VO-enabled tools to
  - select, examine, combine and analyse this dataset, and
  - to combine with *Gaia* photometry and astrometry
Features in Common: data state

• All large surveys will provide reduced data in standardised form, *i.e.* calibrated in standardised units (this will be part of the responsibility of the project itself)

• Data will surely be VO-ready

• Different strategies will probably be employed for data storage
  – databases
  – files
  – VOTables

• Some data will therefore require VO wrapping on access
What Current VO Tools Provide

- Current spectral VO capabilities include
  - VOSpec  Pedro Osuna/ESAC VO Development Team
  - SpecView  Ivo Busko
  - Splat  Peter Draper
  - Spectrum Services for the VO  Tamás Budavári
  ... plus others we will hear about tomorrow

- Tools appear to be generally similar, but with some complementary features, and slightly different approaches

- Tools appear to be orientated mostly towards visualisation, together with some data reduction and analysis capabilities, eg:
  - dereddening of spectra
  - flux measurement
  - fits to generic templates (eg blackbodies) (perhaps to libraries also)
What we might like to do with future large surveys

• Because the data from future large spectroscopic surveys will already be reduced and calibrated, the main thrust of future tools should be in
  – the analysis of large datasets
  – combination of large datasets from different sources
  – data mining
  – visualisations of analysis results, correlations etc.

  … areas in which the VO is the obvious way forward.

• Getting hold of the data within a reasonable amount of time will be important: the scale of future large spectroscopic surveys may drive the methodology of storage, access, processing and delivery of products.

• Intuitiveness of use will be vital if the tools are to be adopted in a wholesale manner by the community.
What we might like to do with future large surveys: scale

• the scale of most future surveys will prohibit the working with and inspection of individual spectra

• some sort of selection and visualisation of outliers will be necessary

• feasibility of single VO tools to deal with the range and scale of tasks?

• automation of procedures – building of workflows – may be the only way forward here (as in AstroGrid Workbench), but
  – this needs to be made less daunting that it currently seems
  – implications for the VO capability tools that plug into the workflow
What we might like to do with future large surveys: analysis (1)

- A clear requirement will be parameter extraction from the spectra:
  - equivalent widths
  - line ratios
  - continuum slopes
  - velocity shifts/redshifts
  - rotations/velocity dispersions
- Classification?
- Searches for non-standard (“interesting”) spectra will be important:
  - emission lines
  - QSO
  - binarity/multiplicity
  - unusual stellar types
What we might like to do with future large surveys: analysis (2)

• It may be difficult to do all this in a standardised way; some may be provided already in the database (eg Gaia), but can VO provide this capability cross database, perhaps adapting methods developed already for single large surveys?

• robust methods will be required for reliable operation on a wide range of spectral data (eg continuum fits to M-star spectra)

• careful handling of observational errors (uncertainties) is vital – generally a more difficult problem than the observational values themselves

• an understanding and recording somehow of the statistical biases (completeness, limits) introduced and propagated by any analysis would be very valuable
What we might like to do with future large surveys: data mining

• A major capability that the VO could provide is the capability to determine the connectedness of different datasets.

• This could be to do with spatial connectedness
  – integral field data (MUSE, KMOS)
  – mosaiced fields
  – phase-space connectedness (momentum, energy)

• Or it could be to do with derived parameter connectedness, eg.
  – age
  – metallicity
  – redshift

• In general will involve combination of spectral, photometric and astrometric data, across the electromagnetic spectrum (cross-matching)
What we might like to do with future large surveys: data mining

• To do this, it would help if the VO could bring in tools to
  – determine the connectedness, and to make selections/cuts
  – visualise the samples
  – provide statistically rigorous outputs

• Again, this may need to be done via workflows tailored to addressing particular investigations
Summary

- Looking forward there will be spectroscopic surveys of >10 million spectra
- These are justified on grounds within the surveys themselves, but science return will be significantly enhanced by the capability to combine and analyse spectroscopic surveys of all types (and including astrometry/photometry)
- Combination of data on such a scale will enable a flowering of new science
- Spectroscopic surveys will need the tools to extract information from the spectra and characterise it in a regularised fashion (line ratios etc.)
- It would be helpful if the VO could examine how this may be done
- The generation of workflows in simple intuitive ways may be the key
- Finally: the VO-provided capability to extract information from the large samples, with rigorous statistical treatment of observational uncertainties, statistical biases, and also with visualisation tools, will be hugely scientifically valuable