Sipping from the Firehouse

Lessons learned from searching for rare sources in large databases

#1 First, keep it simple

Do highly restricted searches for objects that you understand very well before or while targeting the unknown (e.g., Y dwarfs).

Test search algorithms on the brightest targets before tackling the fainter (and often more rewarding and buried) targets

#2 **Use all the resources available to you _before_ going to the telescope**

e.g. DSS, SSS, SDSS, 2MASS, DENIS, **IRSA**

Well-defined, smaller area surveys with substantial ancillary data can be more fruitful and efficient than trying to rely solely on data from one survey.

#3 Make maximal use of automated/queue/training observing

Establish straightforward follow-up programs at multiple facilities, maximizing the use of queue scheduling and robotic observing.

Follow-up observations make great high-school/undergraduate programs.

Even graduate students sometimes have better uses of their time.

#4 **A careful and complete sample is better than twice as many discoveries**

Use simple, reproducible selection criteria

Clearly understand selection effects in the survey data themselves

Simulate selection effects on criteria using a more broadly-defined sample

#5

The human eye is a phenomenal search tool

Aim to include as much human interaction with search algorithms as realistically feasible

e.g. Do you take 6 months to develop the "perfect" rejection code, or 6 months to check all your candidates?

Consider "crowd-sourcing" (e.g., SDSS galaxy classification at http://galaxyzoo.org) - higher processing rates than Cray computers

#6 Be open to the unintended discovery

Serendipity is an inevitable outcome of any search program.

e.g. Brown dwarfs in SDSS; L subdwarfs in 2MASS

Maintain records of "rejects" or "second chances" for later follow-up

#7 Collaborate

You can't do it all yourself. Really.

Your contaminant is someone else's main science - it's time to share.

e.g. The quasar-brown dwarf mash-up

A broad team does good science.

#8 **Publish early, publish often***

If you have a great discovery, don't sit on it waiting for more data - get it out there!

Some of the most highly cited works are discovery papers.

Breakthroughs almost always happen in multiples, so be sure to be the first to drop in

*Advice from my thesis advisor Mike Brown