# What is cross-matching?

When investigating astronomical objects, like active galactic nuclei (AGN), astronomers compare data about those objects from different telescopes at different wavelengths.

This requires *positional cross-matching* to find the closest counterpart within a given radius on the sky.

# Introduction to Catalogues

When we create a catalogue from survey images, we start by extracting the list of sources, starts and galaxies using source finding software. We run through the pixels in an image and find peaks that are statistically significant. Then, we group the surrounding pixels and fit a function, usually based on the telescope response, which is called 'beam' or the 'point spread function'. The result is a list of objects, each of which has a position, an angular size and an intensity measurement. The uncertainties on these measured values depend on things like noise in the image, the calibration of the telescope, and how well we can characterize the telescope's response function. Once we have our catalogues, cross-matching involves searching the second catalogue to find a counterpart for each object in the first catalogue. To do this, we usually search within a given radius, based on the uncertainties in the position. Now, before we dive into the cross-matching algorithm, we first need to understand the notion of distance in space.

# Distance in Space!

We will be calculating the "projected angular distance" between two objects which is not the same as the euclidean distance between two objects. In an astronomical image, one galaxy could be 15 million light-years away while a galaxy next to it can be 100 million light-years away. Just because they appear close in the image doesn't mean they are anywhere close to each other in space. This is a huge challenge in astronomy!

The positions of stars, galaxies and other astronomical objects are usually recorded in either [equatorial](http://astronomy.swin.edu.au/cosmos/E/Equatorial+Coordinate+System) or [Galactic](http://astronomy.swin.edu.au/cosmos/N/North+Galactic+Pole) coordinates.

Equatorial coordinates are fixed relative to the celestial sphere, so the positions are independent of when or where the observations took place. They are defined relative to the celestial equator (which is in the same plane as the Earth's equator) and the ecliptic (the path the sun traces throughout the year).

A point on the celestial sphere is given by two coordinates:

* **Right ascension**: the angle from the vernal equinox to the point, going east along the celestial equator;
* **Declination**: the angle from the celestial equator to the point, going north (negative values indicate going south).

The vernal equinox is the intersection of the celestial equator and the ecliptic where the ecliptic rises above the celestial equator going further east.

Right ascension is often given in **hours-minutes-seconds** (HMS) notation because it was convenient to calculate when a star would appear over the horizon. A full circle in HMS notation is 24 hours, which means 1 hour in HMS notation is equal to 15 degrees. Each hour is split into 60 minutes and each minute into 60 seconds.

Declination, on the other hand, is traditionally recorded in **degrees-minutes-seconds** (DMS) notation. A full circle is 360 degrees, each degree has 60 arcminutes and each arcminute has 60 arcseconds.

To crossmatch two catalogues we need to compare the angular distance between objects on the celestial sphere.

People loosely call this a "distance", but technically it's an *angular* distance: the projected angle between objects as seen from Earth.

If we have an object on the celestial sphere with right ascension and declination (α1,δ1), then the angular distance to another object with coordinates (α2,δ2) is:

d=2arcsin√sin2|δ1−δ2|2+cosδ1cosδ2sin2|α1−α2|2

Angular distances have the same units as angles (degrees). There are [other equations](https://en.wikipedia.org/wiki/Great-circle_distance#Formulas) for calculating the angular distance but this one, called the [haversine formula](https://en.wikipedia.org/wiki/Haversine_formula), is good at avoiding floating-point errors when the two points are close together.

# Cross-Matching Algorithm

for source\_A in catalouge\_1:

for source\_B in catalouge\_2:

offset = angular\_distance(A, B)

if offset < radius:

if offset smallest seen so far:

best\_match = (A, B, offset)