Short Author List Codes

P. Wesley Ryan (ASCL, wes@ascl.net)

Introduction

The literature contains articles on large, community-developed codes, but has comparatively little about the body of software written by 1, 2, or 3 people, what we here refer to as "short author list" (SAL) codes.

What we did

We grouped the codes in the ASCL into three categories: short author list codes, long author list codes, and team/institution codes.

How we did it

Authors in the ASCL are credited in *Lastname, Firstname* format, separated by semicolons, so we counted the commas in ASCL code author list entries to determine the number of individual authors each code has.

For most codes, $n_{semicolons} = n_{commas} - 1$. Some codes list institutions or named collaborations as authors. Since the names of these institutions typically don't contain commas, this equation isn't true for them. This allows us to identify team/institution codes.

What we found out

68% of codes are "short author list" codes, with fewer than four individual authors.

66% of citations to ASCL entries are to short author list codes.

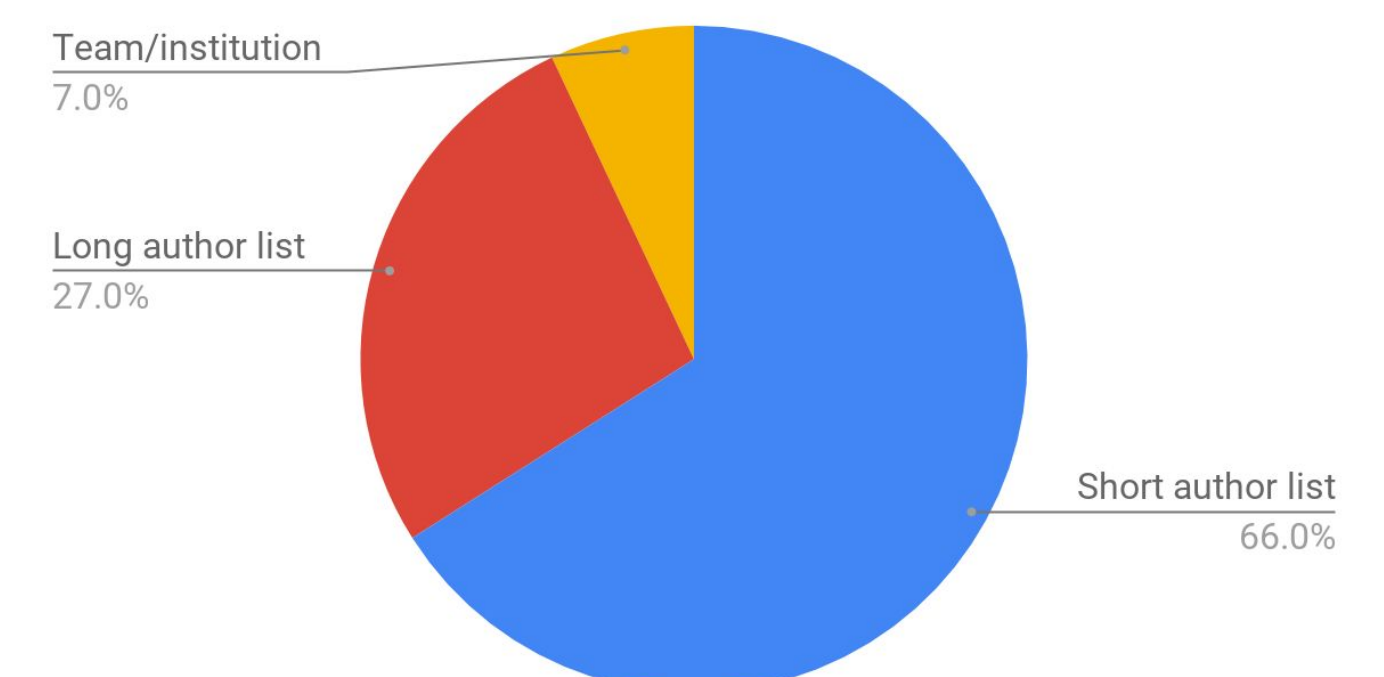Team/institution codes account for only 4% of all codes.

7% of citations to ASCL entries are to team/institution codes.

68% of codes have fewer than 4 authors. 4% of codes are attributed to named teams.

Codes

Team/institution
4.0%

Long author list
27.3%

Short author list
68.7%

Code citations

Team/institution
7.0%

Long author list
27.0%
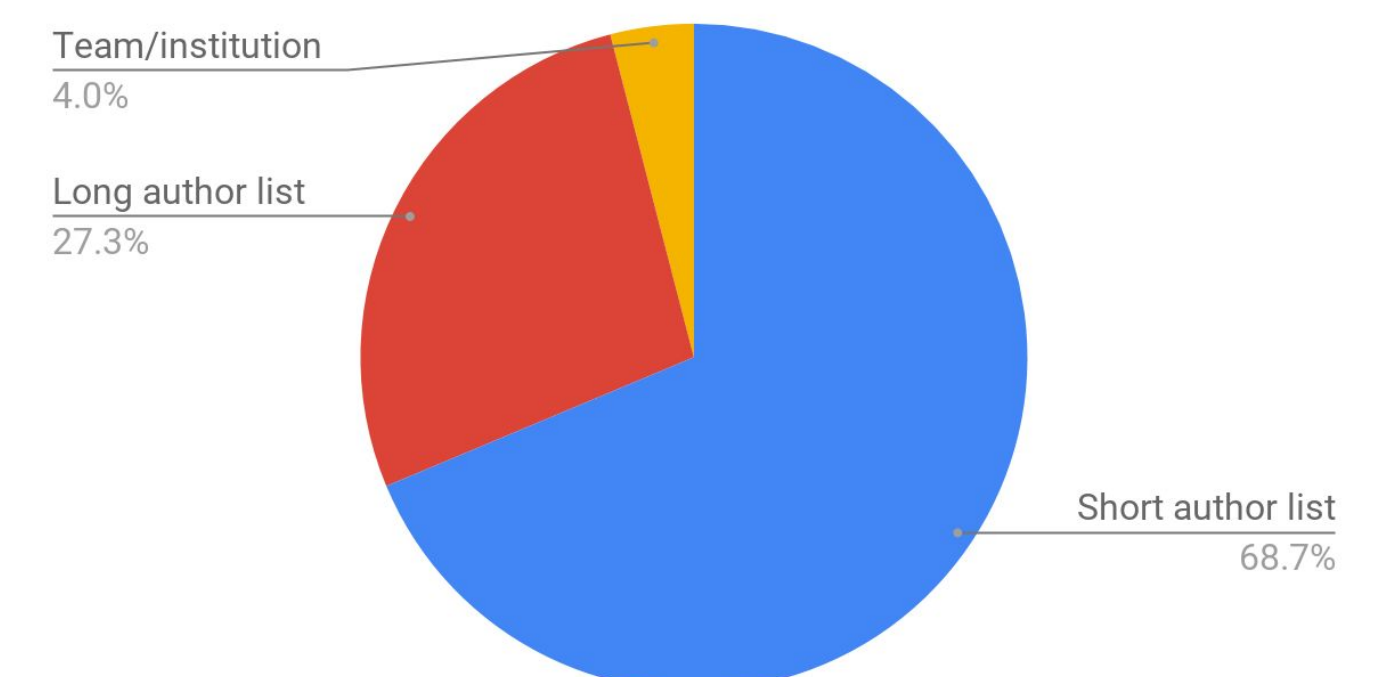
Short author list
66.0%

Conclusions

The majority of ASCL codes have short author lists. Short author list codes have a mean of 110 citations and a median of 22. Team codes have a mean of 207 citations and a median of 27.

Further research

Citations to *described in* or *preferred citation method* papers are not entirely reliable as metrics of code usage. Some of these papers present not only the code in question, but also results obtained with that code; these papers may then be cited only for the results they present.

There are certain cases where codes' *described in* papers are widely cited, but the codes themselves are not widely used, as confirmed by ADS searches for the name of the code.

We are currently investigating the reliability of paper citations as a proxy for code usage, and looking for other metrics of code usage.