Research Tools Completeness Score Report

NF-OSI

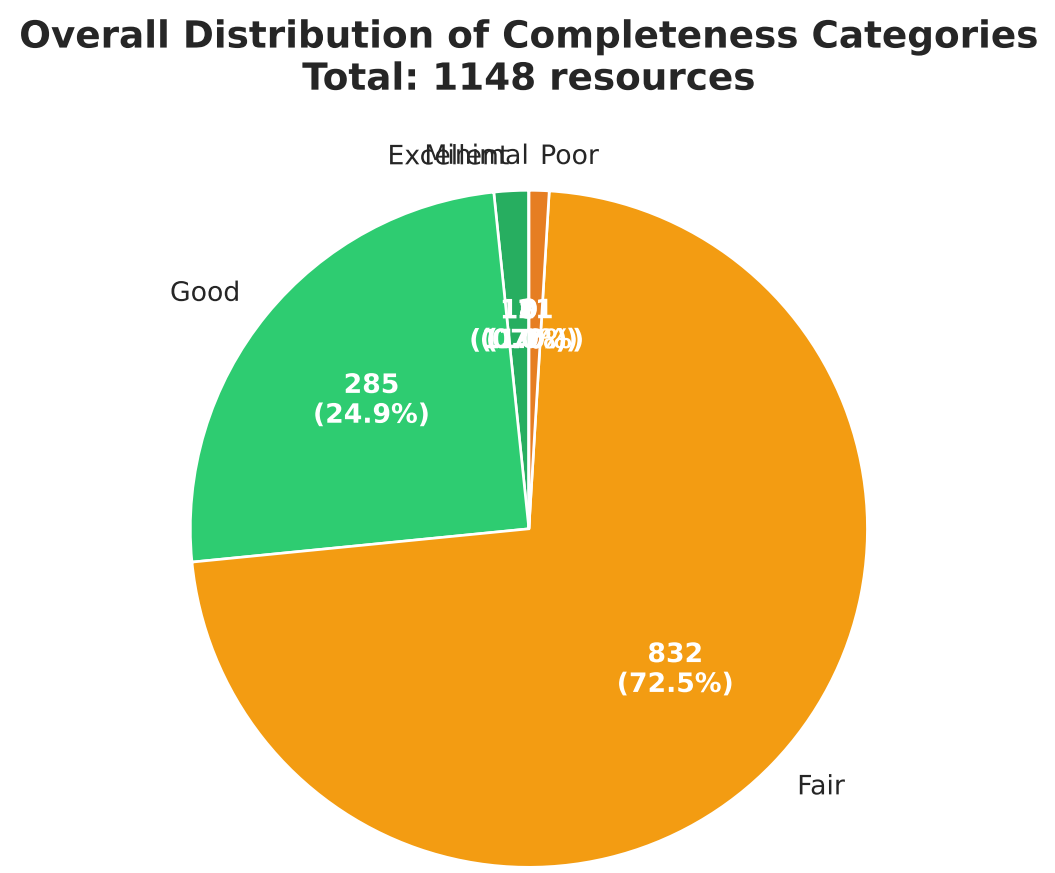
Date: 2025-11-26

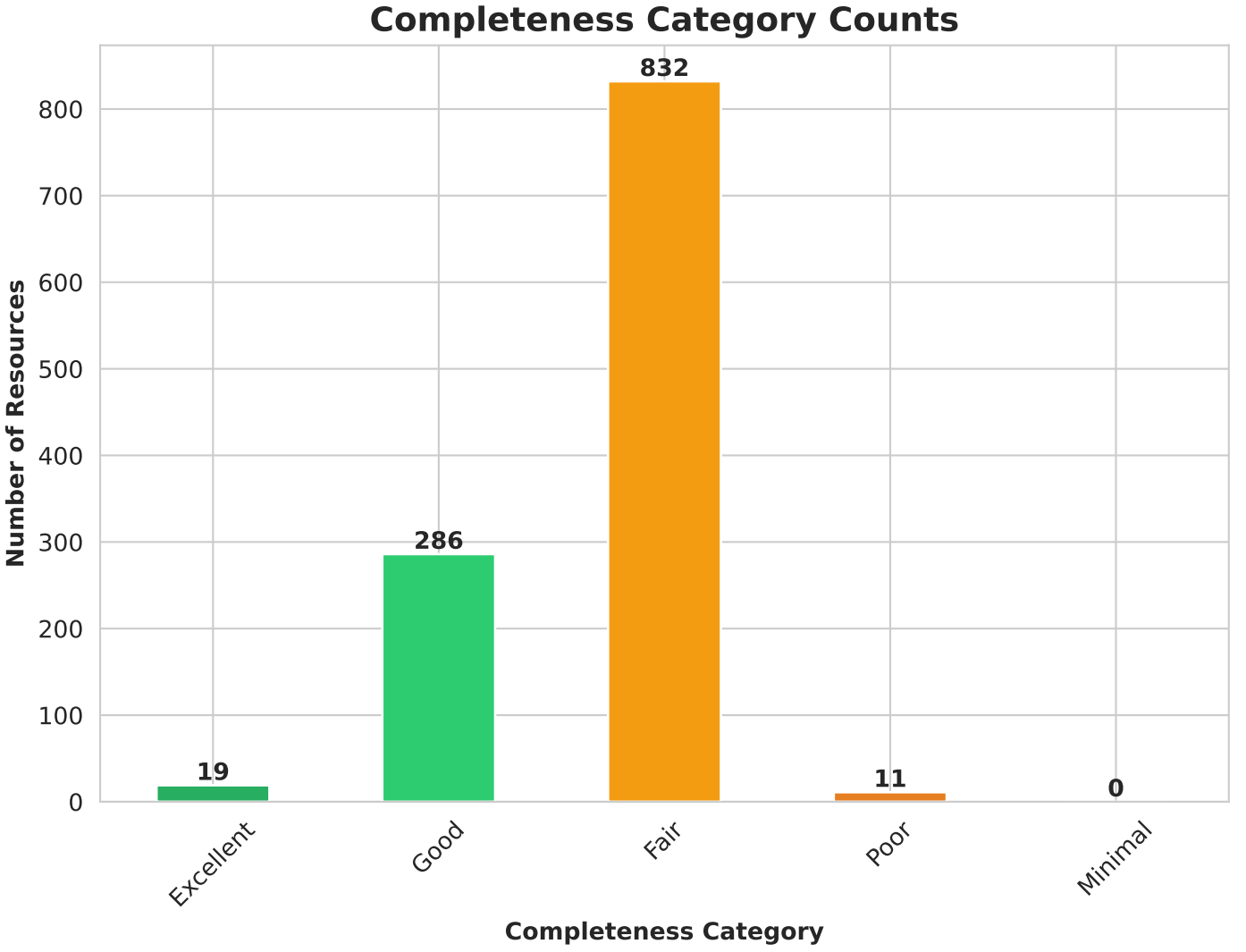
This report provides a comprehensive analysis of the completeness scores for research tools and biobanks in the NF-OSI database.

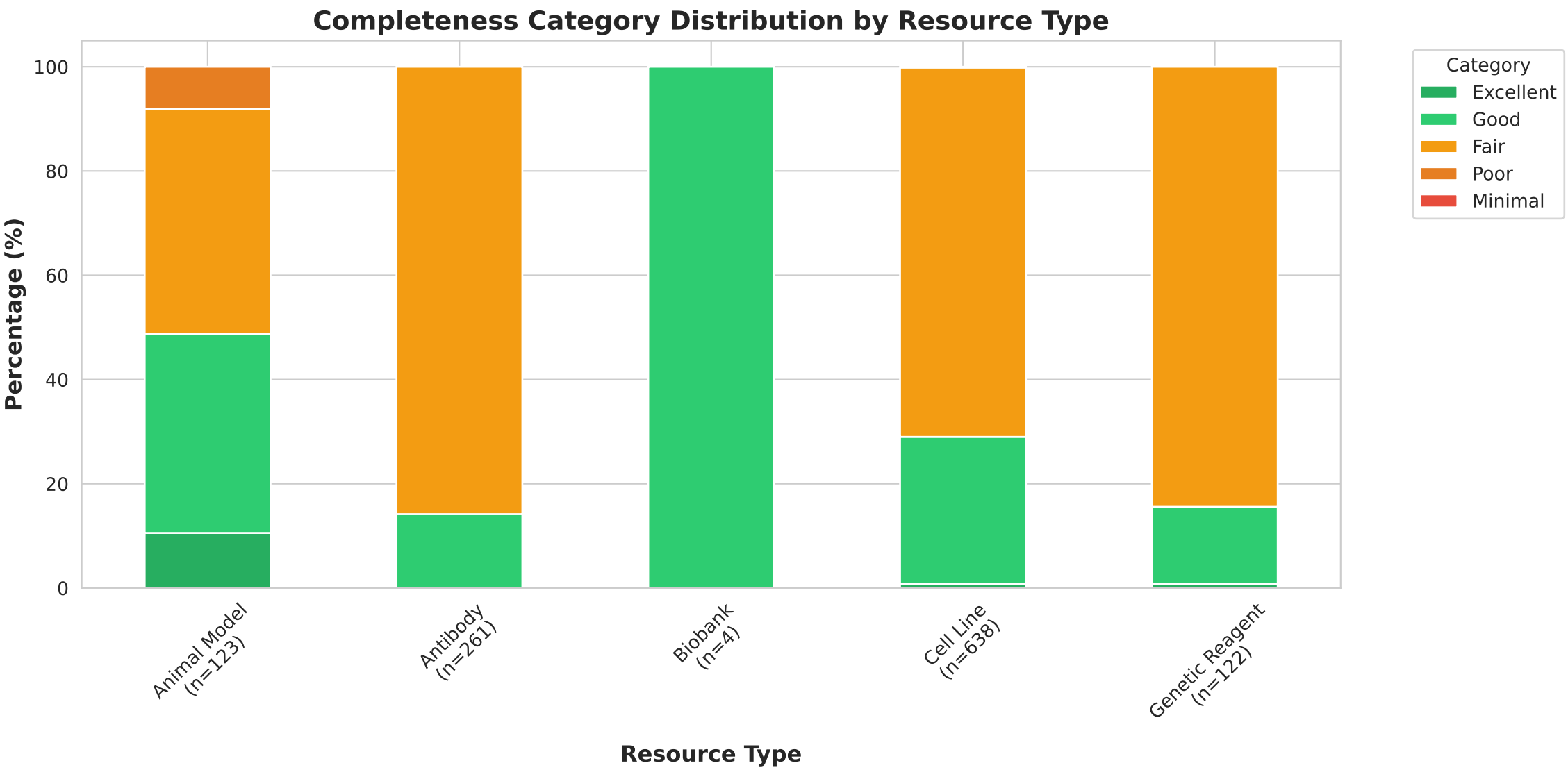
Total Resources Analyzed: 1148
Resource Types: 5

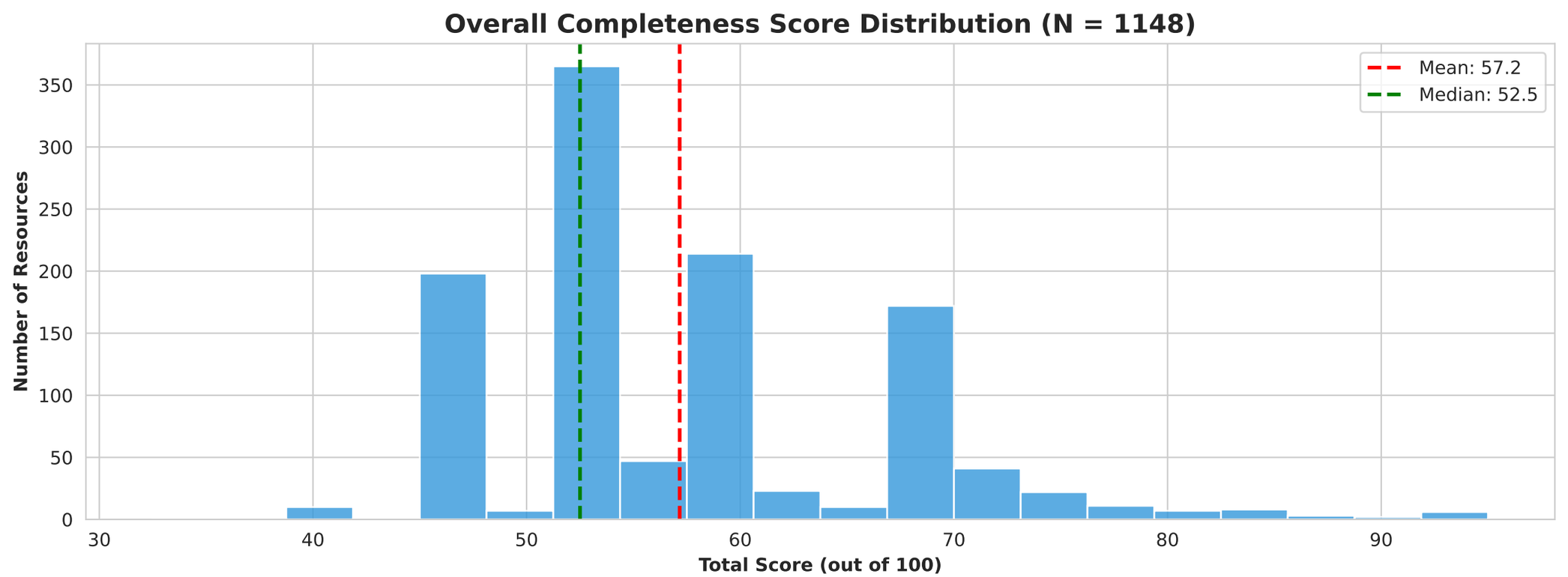
Scoring System (Total: 100 points):
• Availability (30 points): Biobank URL, vendor/developer info, RRID, and DOI
• Critical Info (30 points): Type-specific essential fields
• Other Info (15 points): Type-specific additional fields
• Observations (25 points): Scientific characterizations with DOI weighting

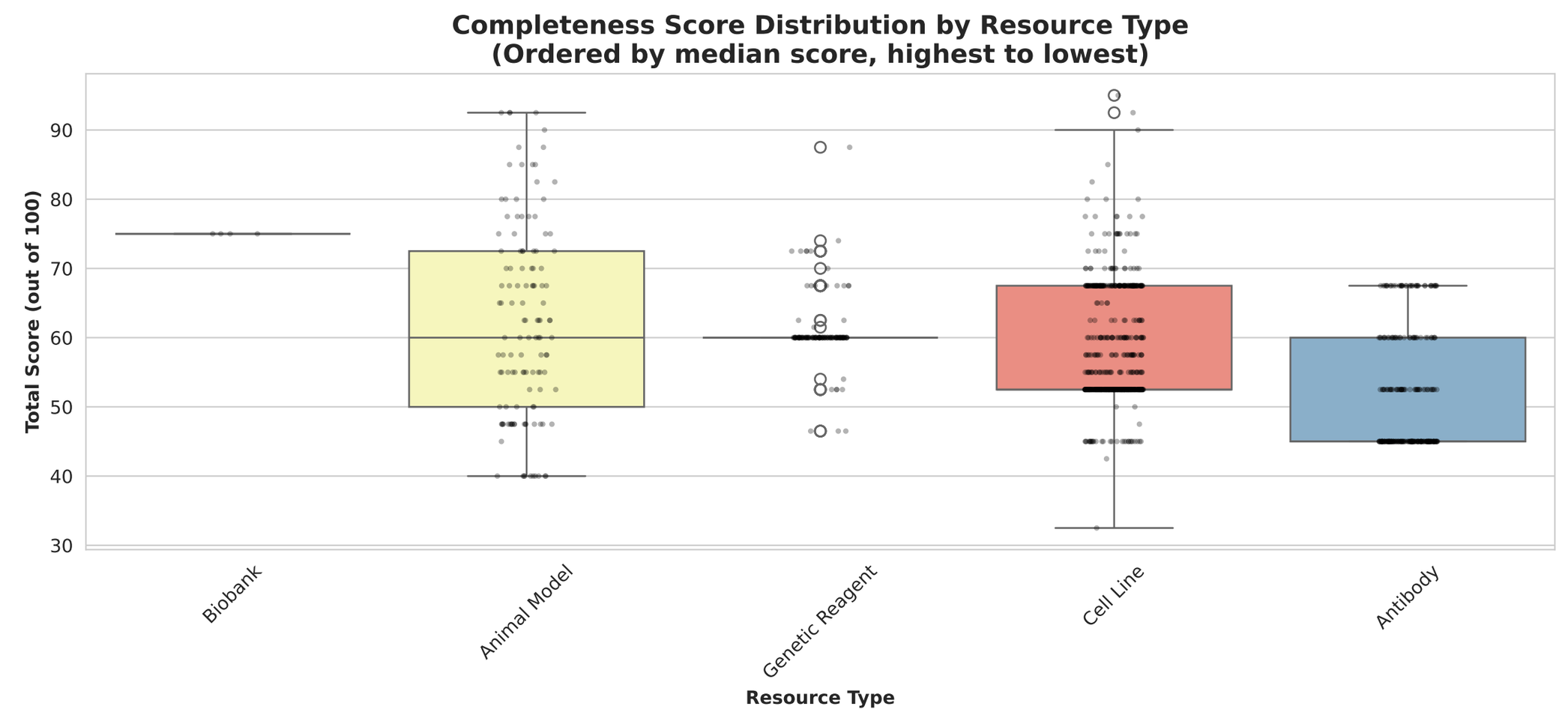
Completeness Categories:
• Excellent: 80+ points
• Good: 60-80 points
• Fair: 40-60 points
• Poor: 20-40 points
• Minimal: <20 points

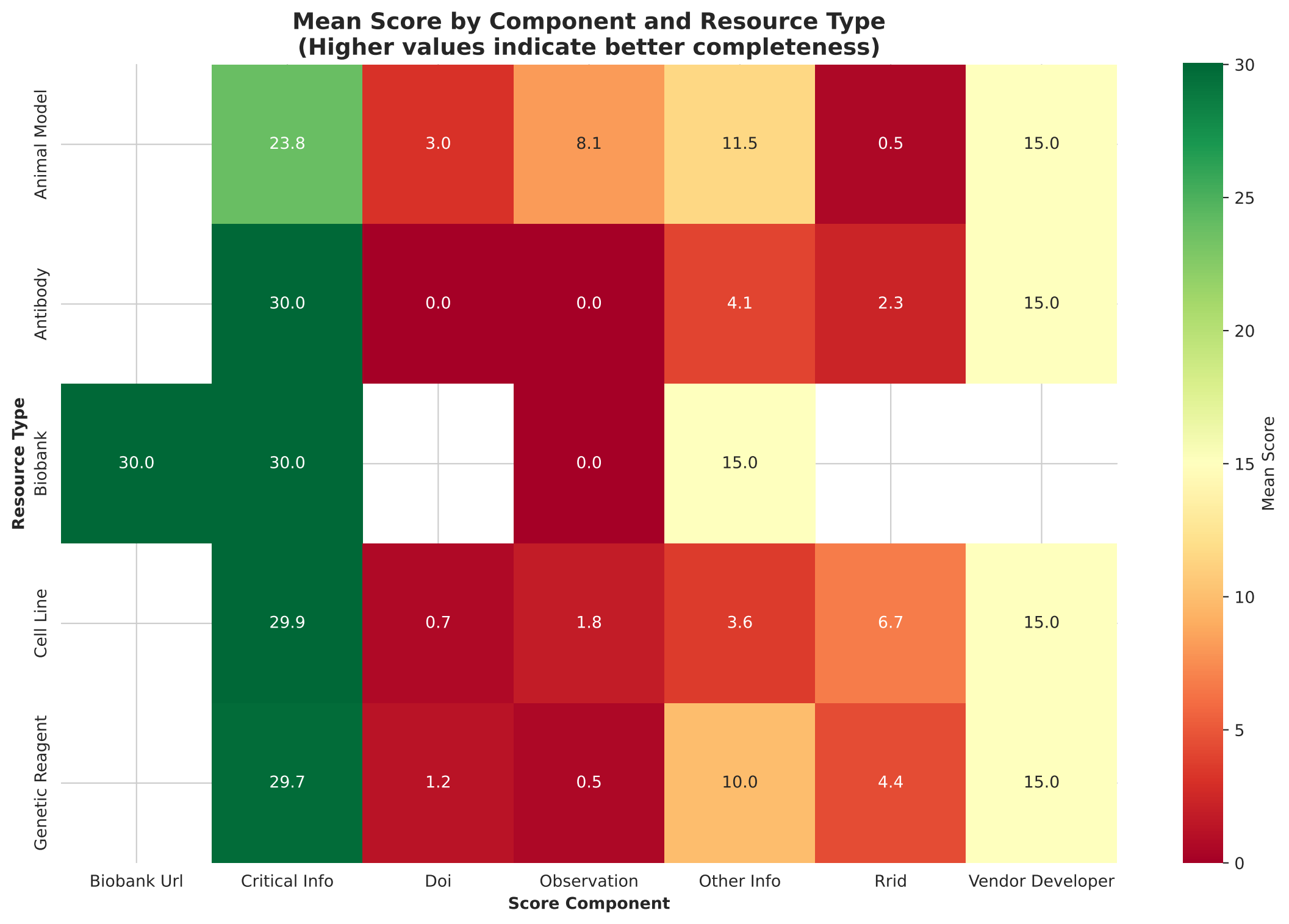

Overall Distribution of Completeness Categories
Total: 1148 resources
Excellent
Minimal
Poor
Good
285
(24.9%)
832
(72.5%)
Fair

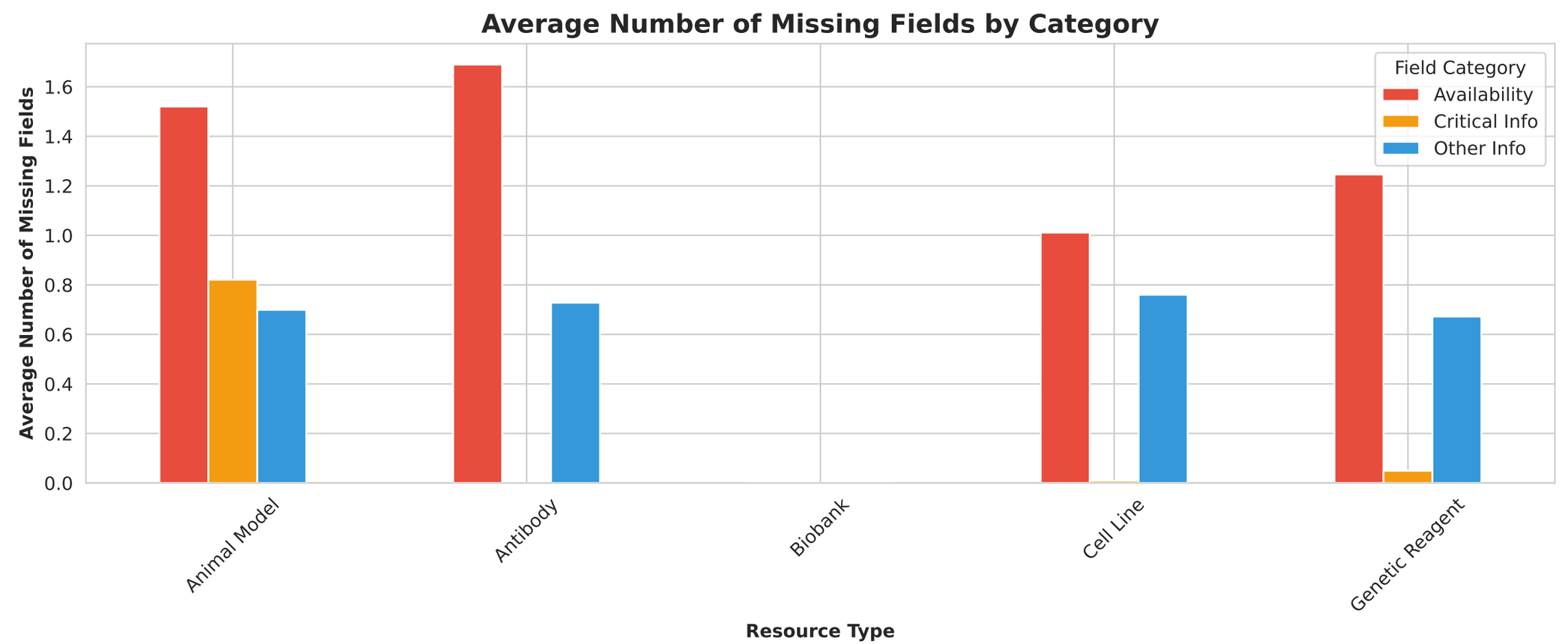

Completeness Category Counts
Number of Resources
0
100
200
300
400
500
600
700
800
19
286
832
11
0
Excellent
Good
Fair
Poor
Minimal
Completeness Category

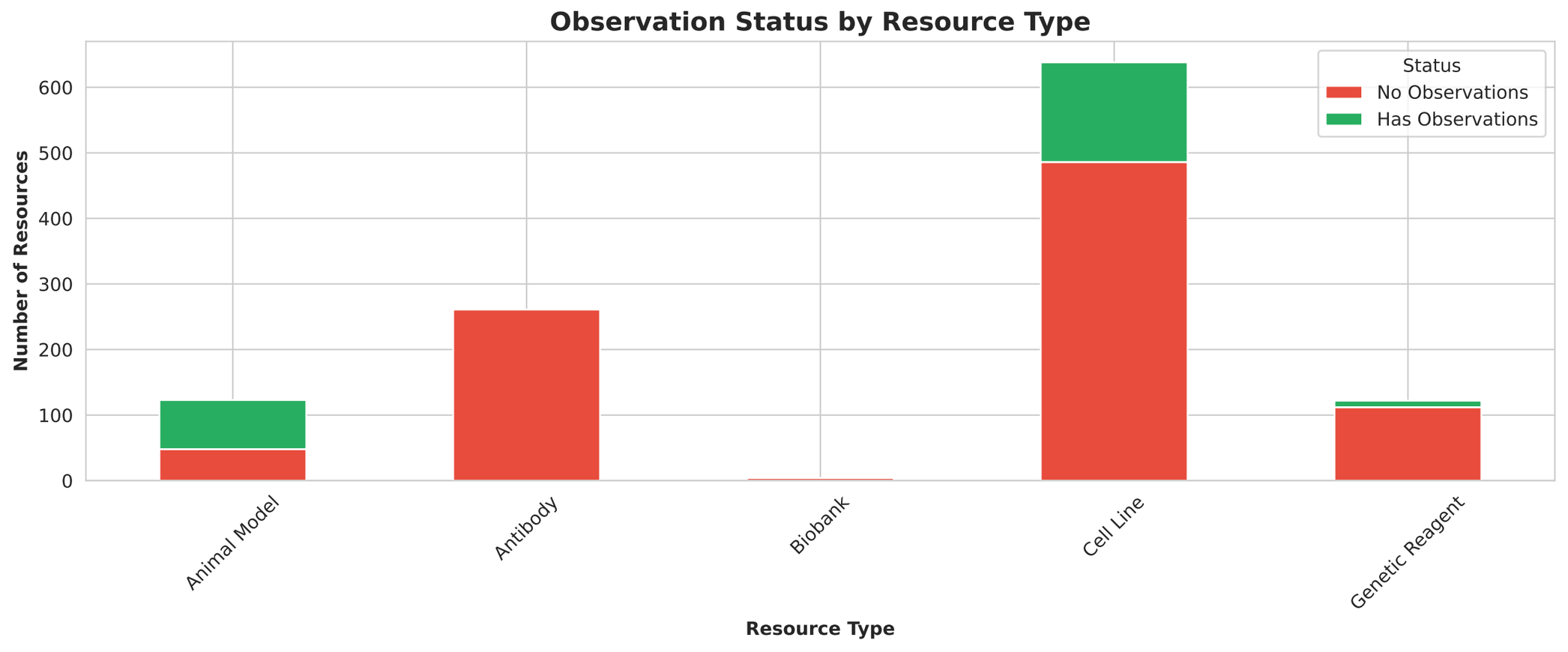

Completeness Category Distribution by Resource Type
Percentage (%)
0
20
40
60
80
100
Animal Model (n=123)
Antibody (n=261)
Biobank (n=4)
Cell Line (n=638)
Genetic Reagent (n=122)
Resource Type
Category
Excellent
Good
Fair
Poor
Minimal


Overall Completeness Score Distribution (N = 1148)
Mean: 57.2
Median: 52.5
Number of Resources
Total Score (out of 100)
0
50
100
150
200
250
300
350
30
40
50
60
70
80
90


Completeness Score Distribution by Resource Type
(Ordered by median score, highest to lowest)
Total Score (out of 100)
30
40
50
60
70
80
90
Biobank
Animal Model
Genetic Reagent
Cell Line
Antibody
Resource Type


Mean Score by Component and Resource Type
(Higher values indicate better completeness)
Resource Type
Animal Model
Antibody
Biobank
Cell Line
Genetic Reagent
23.8
3.0
8.1
11.5
0.5
15.0
30.0
0.0
0.0
4.1
2.3
15.0
30.0
30.0
0.0
15.0
29.9
0.7
1.8
3.6
6.7
15.0
29.7
1.2
0.5
10.0
4.4
15.0
Biobank Url
Critical Info
Doi
Observation
Other Info
Rrid
Vendor Developer
Score Component
Mean Score
30
25
20
15
10
5
0


Average Number of Missing Fields by Category
Average Number of Missing Fields
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
Field Category
Availability
Critical Info
Other Info
Animal Model
Antibody
Biobank
Cell Line
Genetic Reagent
Resource Type


Observation Status by Resource Type
Number of Resources
0
100
200
300
400
500
600
Status
No Observations
Has Observations
Animal Model
Antibody
Biobank
Cell Line
Genetic Reagent
Resource Type

Overall Summary Statistics

Metric	Value
Total Resources	1148.0
Mean Score	57.16
Median Score	52.5
Std Dev	9.48
Min Score	32.5
Max Score	95.0
Range	62.5

Summary Statistics by Resource Type

Resource Type	Count	Mean	Median	Std Dev	Min	Max
Biobank	4	75.0	75.0	0.0	75.0	75.0
Animal Model	123	62.07	60.0	14.05	40.0	92.5
Genetic Reagent	122	60.85	60.0	4.99	46.5	87.5
Cell Line	638	57.75	52.5	8.22	32.5	95.0
Antibody	261	51.41	45.0	8.35	45.0	67.5

Resources with 10 Lowest Scores

Resource Name	Type	RRID	Score	Category
HS-PSS	Cell Line	rrid:CVCL_8717	32.5	Poor
Krox20;Nf1flox/-	Animal Model	N/A	40.0	Poor
miR-155 +/-;Nf1flox/+	Animal Model	N/A	40.0	Poor
MPNST-Nf1-001	Animal Model	N/A	40.0	Poor
MPNST-Nf1-002	Animal Model	N/A	40.0	Poor
MPNST-SP-001	Animal Model	N/A	40.0	Poor
MPSNT-SP-002	Animal Model	N/A	40.0	Poor
Nf1 GEM	Animal Model	N/A	40.0	Poor
Nf1-/- Drosophila	Animal Model	N/A	40.0	Poor
Nf1Fcr	Animal Model	N/A	40.0	Poor

Recommendations

Based on the analysis, here are key recommendations for improving resource completeness:

1. Focus on Low-Scoring Components
   Identify which score components (availability, critical info, other info, observations) have the lowest average scores across resource types and prioritize filling those fields.

2. Resource Type Priorities
   Target resource types with lower median scores for systematic improvement efforts. Review the boxplot and summary statistics to identify priorities.

3. Observation Documentation
   Encourage researchers to submit observations with publication DOIs, as these contribute more points (7.5 vs 2.5) to the completeness score. Currently, many resources lack observations entirely.

4. RRID Registration
   Ensure all resources have registered RRIDs (Research Resource Identifiers) to improve findability and citation. This is worth 7.5 points in the availability category.

5. Vendor/Developer Information
   Complete vendor or developer information for resources to improve availability scores. This accounts for 15 points in the scoring system.

6. Address Low-Scoring Resources
   Prioritize improving these to reach at least "Fair" completeness status.

7. Type-Specific Fields
   Review the critical info and other info fields specific to each resource type. These account for 45 points total and are often incomplete.

8. Regular Monitoring
   Schedule regular reviews of completeness scores to track improvement over time and identify new resources that need attention.