PART 2 – Cross Dataset Engineering Tasks

Task 12: Material Descriptor Analysis

May 1, 2025

Abstract

This report presents a detailed analysis of engineering materials based on Task 12 (Material Descriptor Analysis) of the materials dataset. We examine the relationship between descriptive text data and material properties, focusing on keywords like "heat-resistant", "heat-treatment", and "corrosion-resistant". The analysis provides critical insights into material selection criteria based on text descriptions, revealing unique property profiles for different material categories and identifying specialized materials that combine multiple desirable properties. Data quality observations are also included to inform the reliability of engineering decisions based on text-based classification.

1 Introduction

The Value of Unstructured Text Data in Materials Selection

Understanding how text descriptors relate to material properties is increasingly important in modern engineering materials selection. This report analyzes the dataset by extracting insights from the Desc column, identifying keyword patterns, and correlating them with mechanical, physical, and chemical properties. By examining how property profiles vary across different descriptor categories, engineers can leverage text-based classification to make more informed material selection decisions, particularly for specialized applications requiring specific combinations of properties.

2 Material Selection Insights

2.1 Keyword Distribution Analysis

The analysis of descriptive keywords reveals significant patterns in the frequency and distribution of material categories:

Table 1: Material Distribution by Descriptor Keyword

Keyword Category	Material Count	Prevalence
Heat-treatment	132	Dominant category (56.4%)
Heat-resistant	66	Moderate presence (28.2%)
Corrosion-resistant	8	Specialized category (3.4%)
Others/Uncategorized	28	Miscellaneous (12.0%)

Key Insights from Keyword Distribution

The distribution analysis of descriptive terms reveals important patterns for engineering material classification:

- **Category Dominance**: Heat-treatment materials form the largest group (132 instances), indicating the prevalence of thermally processed materials in engineering applications.
- **Specialized Nature**: Corrosion-resistant materials form a small, specialized category (only 8 instances), suggesting these materials are more specialized and potentially higher-value.
- **Data Coverage**: Of 1552 original materials, only 981 have descriptive text data (63.2% coverage), limiting the comprehensiveness of text-based classification.
- **Descriptor Diversity**: The dataset contains 83 unique descriptor values, indicating a lack of standardized terminology that could impact classification consistency.

2.2 Property Profile by Material Category

Different descriptor categories exhibit distinct property profiles, creating specific performance characteristics:

Table 2: Average Material Properties by Descriptor Category

Category	Su (MPa)	Sy (MPa)	A5 (%)	E (MPa)	G (MPa)	mu	Ro (kg/m
Heat-treatment	807.8	589.3	14.7	206,000	80,000	0.300	7,860
Heat-resistant	589.3	345.8	24.1	205,667	79,803	0.300	7,862
Corrosion-resistant	775.0	476.0	25.5	203,250	78,375	0.299	7,874

Property Profile Insights

The analysis of property variations across descriptor categories reveals critical patterns for materials selection:

- **Strength Leadership**: Heat-treatment materials demonstrate superior ultimate tensile strength (Su = 807.8 MPa) and yield strength (Sy = 589.3 MPa), making them optimal for load-bearing applications.
- **Ductility Advantage**: Both heat-resistant and corrosion-resistant materials exhibit significantly higher elongation values (A5 = 24.1% and 25.5% respectively) compared to heat-treatment materials (A5 = 14.7%), making them better suited for applications requiring formability.
- **Modulus Consistency**: Young's modulus (E) and shear modulus (G) remain relatively consistent across all categories, indicating similar elastic behavior regardless of material category.
- **Density Variation**: Corrosion-resistant materials have slightly higher density (Ro = 7,874 kg/m^3) compared to other categories, potentially influencing weight-critical applications.

2.3 Multi-Property Materials Analysis

Analysis of material descriptor overlaps identifies specialized materials with multiple desirable properties:

Table 3: Special Materials with Both Heat and Corrosion Resistance

Material	Key Properties	Application Suitability
Steel 13Ch14N3V2FR GOST 5949-75	High Su, High Sy, Strong elasticity	Turbine disks, blades (high-load, high-temp)
Steel 10Ch11N23T3MR GOST 5949-75	Moderate-High strength, High density	Steam lines, corrosive + thermal environments
Steel 12Ch18N10T GOST 5949-75	Low strength, Very high A5, Superior E	Chemical vessels, food-grade equipment
Steel 12Ch18N9T GOST 5949-75	Slightly higher strength than 10T	Heat exchangers, moderate-load applications
Steel 12Ch18N12T GOST 5949-75	Balanced profile similar to 9T	Cryogenic vessels, acid piping
Steel 12Ch25N16G7AR GOST 5949-75	Moderate strength, Very high A5, Best E	Acidic reactors, marine exhausts

Engineering Insights on Multi-Property Materials

The overlap analysis identifies critical materials with combined properties:

- **Material Classification**: All six materials with both heat and corrosion resistance are Soviet/Russian austenitic stainless steels standardized under GOST 5949-75.
- **Alloying Elements**: These materials contain high levels of chromium (Ch), nickel (N), and specialized elements like titanium (T), vanadium (V), and boron (R), contributing to their exceptional property combinations.
- **Property Range**: Within this elite group, materials like Steel 13Ch14N3V2FR offer superior strength, while others like 12Ch18N10T provide exceptional ductility—enabling application-specific selection.
- **Heat Treatment Sensitivity**: Several materials (e.g., 13Ch14N3V2FR, 10Ch11N23T3MR) appear with multiple property profiles, indicating their properties can be tuned through different heat treatments.

3 Application-Specific Recommendations

Based on the material descriptor analysis, specific materials can be recommended for various engineering applications:

Table 4: Application Recommendations Based on Material Descriptors

Application	Recommended Category	Specific Material Examples
High-load, high temperature	Heat+Corrosion Resistant	Steel 13Ch14N3V2FR GOST 5949-75
Chemical processing vessels	Corrosion Resistant	Steel 12Ch18N10T GOST 5949-75
Power generation components	Heat Resistant	Steel 10Ch11N23T3MR GOST 5949-75
Structural components	Heat Treatment	Heat-treatment steel variants
Marine applications	Corrosion Resistant	Steel 12Ch25N16G7AR GOST 5949-75
Aerospace components	Balanced properties	Steel 12Ch18N12T GOST 5949-75

Application-Specific Selection Criteria

The descriptor analysis supports application-specific material selection through these insights:

- **Environmental Challenge**: For combined thermal and corrosive environments (e.g., chemical processing, power generation), multi-property materials from the overlapping category are ideal choices.

- **Property Prioritization**: Applications can be matched with materials based on their descriptor-associated property profiles—strength-dominated (heat-treatment), ductility-dominated (heat-resistant), or balance-focused (corrosion-resistant).
- **Specialized Applications**: For critical applications like aerospace components or medical devices, materials with carefully balanced properties like Steel 12Ch18N12T offer optimal performance.
- **Cost-Performance Balance**: For less demanding applications, standard heat-treatment materials provide adequate performance at potentially lower cost than specialized heat-resistant or corrosion-resistant alternatives.

4 Data Quality Observations

4.1 Descriptor Coverage and Standardization

The descriptor analysis reveals several data quality issues affecting the comprehensiveness and reliability of text-based classification:

Table 5: Data Quality Issues in Material Descriptors

Issue	Observation	Impact on Analysis
Descriptor Coverage	36.8% of materials lack descriptors	Limited comprehensiveness of text analysis
Descriptor Standardization	83 unique descriptor values present	Potential inconsistency in classification
Property Completeness	Bhn values missing for many materials	Incomplete property comparisons
Heat Treatment Variation	Multiple entries for same material	Complicates material identification

4.2 Material Property Variations and Consistency

Property Measurement and Reporting Issues

Several observations regarding material property data quality emerged from the descriptor analysis:

- **Missing Hardness Data**: Brinell Hardness (Bhn) values appear to be missing or insignificant across analyzed descriptor categories, limiting the comparison of surface mechanical properties.

- **Property Variation with Heat Treatment**: Materials like Steel 13Ch14N3V2FR show property variations with different heat treatments, introducing multiple data points for single materials and complicating classification.
- **Descriptor Inconsistency**: The lack of standardized terminology (83 unique descriptors) suggests potential for classification errors or inconsistencies in material grouping.
- **Missing Materials**: With 36.8% of materials lacking descriptors, certain material categories may be underrepresented in the analysis, potentially skewing the observed patterns.

5 Engineering Implications

The material descriptor analysis provides valuable guidance for engineering design and material selection decisions:

Practical Applications of Text-Based Material Selection

Engineers can leverage text descriptors in several practical ways:

- **Selection Trade-offs**: The clear strength-ductility trade-off observed between heat-treatment materials (higher strength, lower ductility) and heat/corrosion-resistant materials (lower strength, higher ductility) provides a framework for balancing competing requirements.
- **Material Tunability**: The observation that several materials exhibit property variations with different heat treatments demonstrates the potential for fine-tuning material performance to specific operational conditions.
- **Specialized Applications**: The identification of rare materials with both heat and corrosion resistance provides valuable options for extreme environments, though these should be reserved for critical applications due to likely higher cost.
- **Text-Based Pre-screening**: Descriptive text provides a valuable first-pass filter for material selection, enabling engineers to quickly narrow the field of candidates before detailed property comparison.

6 Conclusions and Recommendations

The material descriptor analysis yields several key conclusions for engineering material selection:

Key Descriptor Analysis Takeaways

The analysis of materials by descriptive text reveals:

- Descriptive terminology provides valuable classification information that complements numerical properties, demonstrating the importance of unstructured data in technical domains.
- Heat-treatment materials offer superior strength properties, heat-resistant materials provide enhanced ductility, and corrosion-resistant materials deliver specialized environmental performance—creating distinct selection categories.
- Six special austenitic stainless steels combine both heat resistance and corrosion resistance, providing elite options for extreme environments requiring multiple property advantages.
- Data quality issues, particularly missing descriptors and hardness values, require careful consideration when making engineering decisions based on text categorization.
- Text-based material selection offers a complementary approach to traditional property-based selection, potentially streamlining the identification of materials for specialized applications.

6.1 Future Analysis Recommendations

To extend the value of material descriptor analysis, several future analytical approaches are recommended:

Advancing Text-Based Materials Analysis

Further analysis could enhance the value of material descriptors through:

- **Expanded Keyword Analysis**: Extend the analysis to additional descriptive terms beyond the three primary keywords to identify other specialized material categories.
- **Correlation Analysis**: Perform statistical analysis to quantify the relationships between specific descriptive terms and material properties, enabling more precise prediction of performance.
- **Natural Language Processing**: Apply more sophisticated text analysis techniques to extract patterns from unstructured material descriptions, potentially uncovering hidden relationships.
- **Descriptor Standardization**: Develop a standardized classification system for material descriptions to improve consistency and comparability across the dataset.

- **Property Gap Analysis**: Identify specific properties missing from different descriptor categories to guide future data collection efforts and improve analysis completeness.

7 Appendix: Material Descriptor Analysis Methodology and PLots

Analysis Process

The material descriptor analysis followed these steps:

1. **Data Preparation**: Filtering the dataset to include only rows with non-null descriptor values (981 of 1552 materials).
2. **Keyword Identification**: Selecting three primary keywords for analysis: "heat-resistant", "heat-treatment", and "corrosion-resistant".
3. **Category Formation**: Creating material categories based on the presence of these keywords in the descriptor field.
4. **Property Aggregation**: Calculating mean values for mechanical and physical properties (Su, Sy, A5, E, G, mu, Ro) within each category.
5. **Overlap Analysis**: Identifying materials belonging to multiple keyword categories to find multi-property materials.
6. **Visualization**: Creating radar charts and Venn diagrams to visualize property differences and category relationships.
7. **Application Mapping**: Connecting material categories and their property profiles with engineering application requirements.

Radar Chart Insights Summary

The following table summarizes which material group "wins" in each property, based on radar chart median comparisons across groups (e.g., heat-treated, corrosion-resistant).

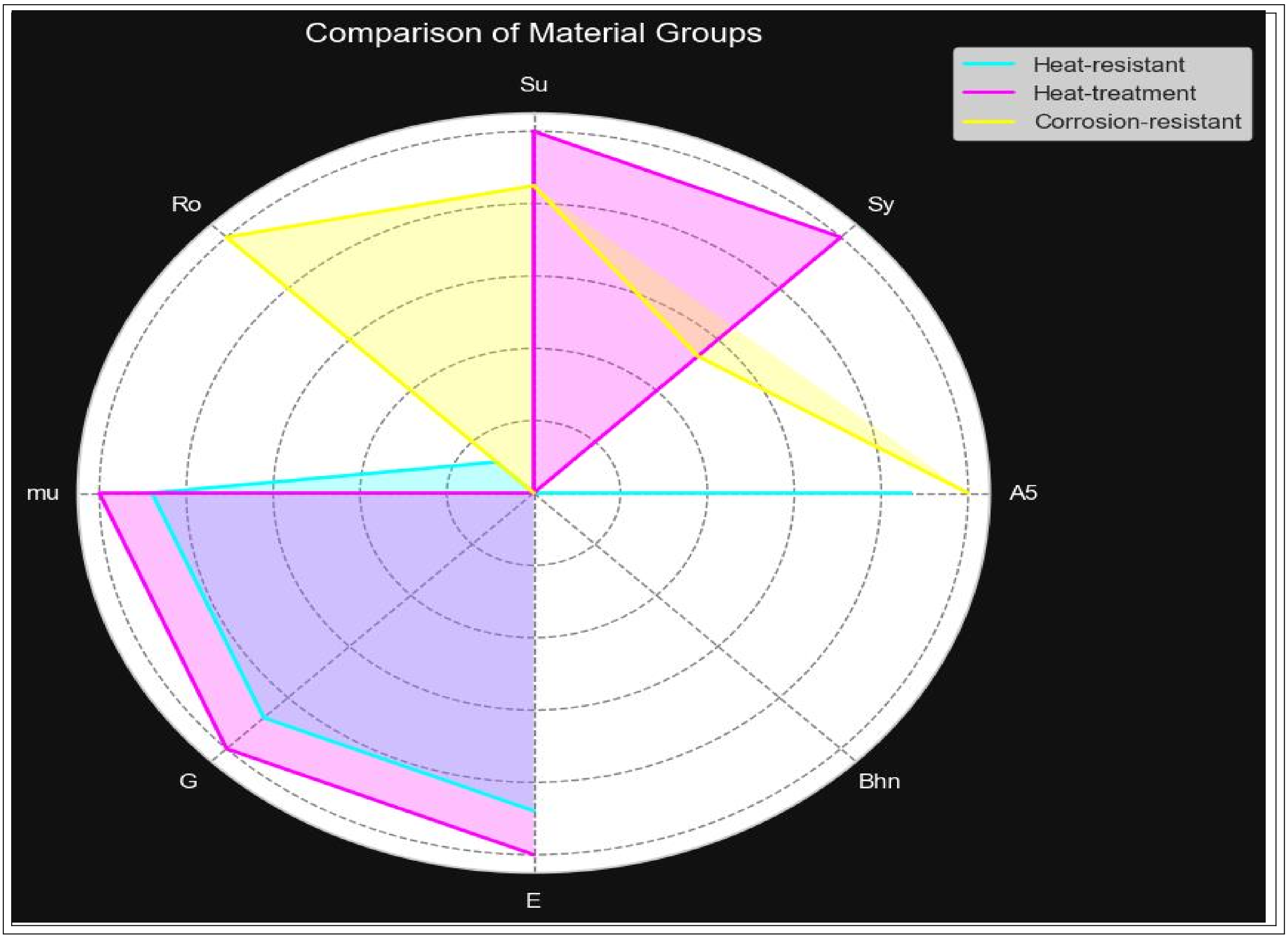


Figure 1: Comparision of Material Groups

Property	Who Wins?	Comment
Su (Ultimate Strength)	Heat-treated (magenta)	Strongest — clear Su advantage after heat treatment
Sy (Yield Strength)	Heat-treated (magenta)	Again dominant in yield strength, showing great plastic resistance
A5 (Elongation)	Corrosion-resistant (yellow)	Most ductile — suitable for flexible/deformable applications
Bhn (Brinell Hardness)	None	Data near zero or missing — not useful in current chart
E (Young's Modulus)	Heat-treated (magenta)	Highest stiffness — resists elastic deformation better
G (Shear Modulus)	Heat-treated (magenta)	Strong shear resistance — good for torsional loads
μ **(Poisson's Ratio)**	Heat-treated (magenta)	Slightly better transverse strain behavior than others
Ro (Density)	Corrosion-resistant (yellow)	Heaviest group — higher density materials dominate

Venn Diagram Interpretation & Overlap Insights

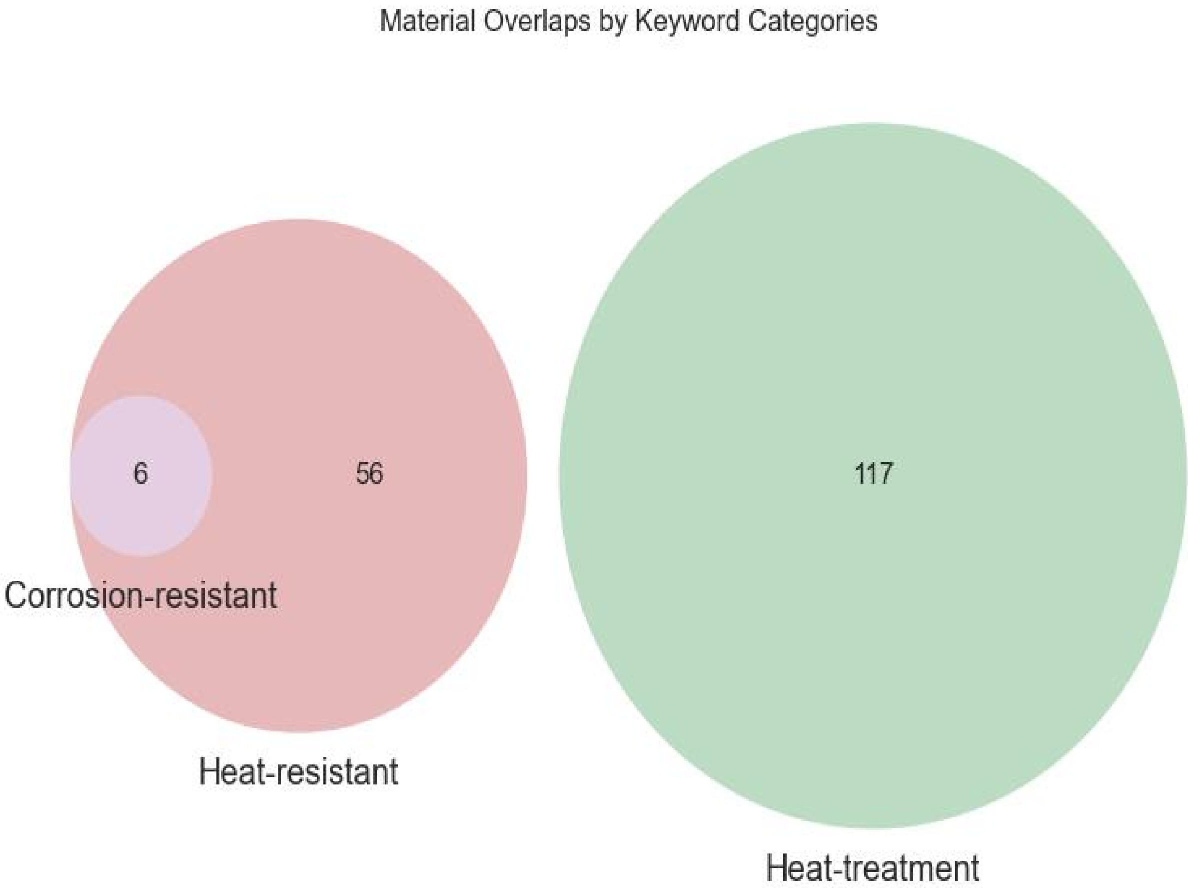


Figure 2: Material Overlap by Keyword Categories

Group Size Summary:

- **Heat-treatment steels:** Largest group (117 materials).
- **Heat-resistant steels:** Mid-size group (56 materials).
- **Corrosion-resistant steels:** Smallest group (6 materials).

Overlap Behavior:

- 6 materials are **both** heat- and corrosion-resistant.
- **Heat-treatment** group is mostly separate — no significant overlap.
- **Corrosion-resistant steels** are a tiny subset of heat-resistant steels.

Engineering Meaning:
Most corrosion-resistant steels are also made to resist heat — a known characteristic of **austenitic stainless steels**, especially in high-performance applications.

Overlapping Steels: Heat & Corrosion Resistant

These 6 materials appear in both categories:

- Steel 10Ch11N23T3MR GOST 5949-75
- Steel 12Ch18N10T GOST 5949-75
- Steel 12Ch18N12T GOST 5949-75
- Steel 12Ch18N9T GOST 5949-75
- Steel 12Ch25N16G7AR GOST 5949-75
- Steel 13Ch14N3V2FR GOST 5949-75

Alloy Symbols:
Ch = Cr, **N = Ni**, **T = Ti**, **G = Mn**, **R = B**, **V = V**
These steels are part of the Soviet/Russian GOST 5949-75 standard and are comparable to Western stainless grades like **AISI 321, 316Ti, 347**.

Engineering Properties to Expect

Property	Behavior
Strength	Excellent at elevated temperatures
Ductility	Good — retains toughness at low temperatures
Corrosion Resistance	Excellent — withstands acids, salt spray, chemicals
Heat Resistance	Very good — due to Cr and Ni content
Weldability	Generally good — Ti used for stabilization
Hardening	Not hardenable by heat — only by cold working

Engineering Summary:
If your application demands **both corrosion and heat resistance**, these 6 steels are top-tier candidates. Added elements like Ti, V, and B enhance:

- Intergranular corrosion resistance
- High-temperature creep strength
- Low-temperature impact toughness

Material-by-Material Engineering Insights

Steel 13Ch14N3V2FR: Excellent Su and Sy, average elasticity (E/G), high Poisson's ratio.
Use Case: Turbine disks, high-load/high-temp components.

Steel 10Ch11N23T3MR: Moderate-to-high strength, balanced elasticity, heavier density.
Use Case: Steam lines, thermal + corrosive environments.

Steel 12Ch18N10T: Low strength but extremely ductile. Very high E and G.
Use Case: Chemical tanks, corrosion-focused environments.

Steel 12Ch18N9T: Slightly stronger than 10T, also highly ductile.
Use Case: Food-grade equipment, exchangers.

Steel 12Ch18N12T: Matches 9T, good balanced profile.
Use Case: Acidic lines, cryogenic vessels.

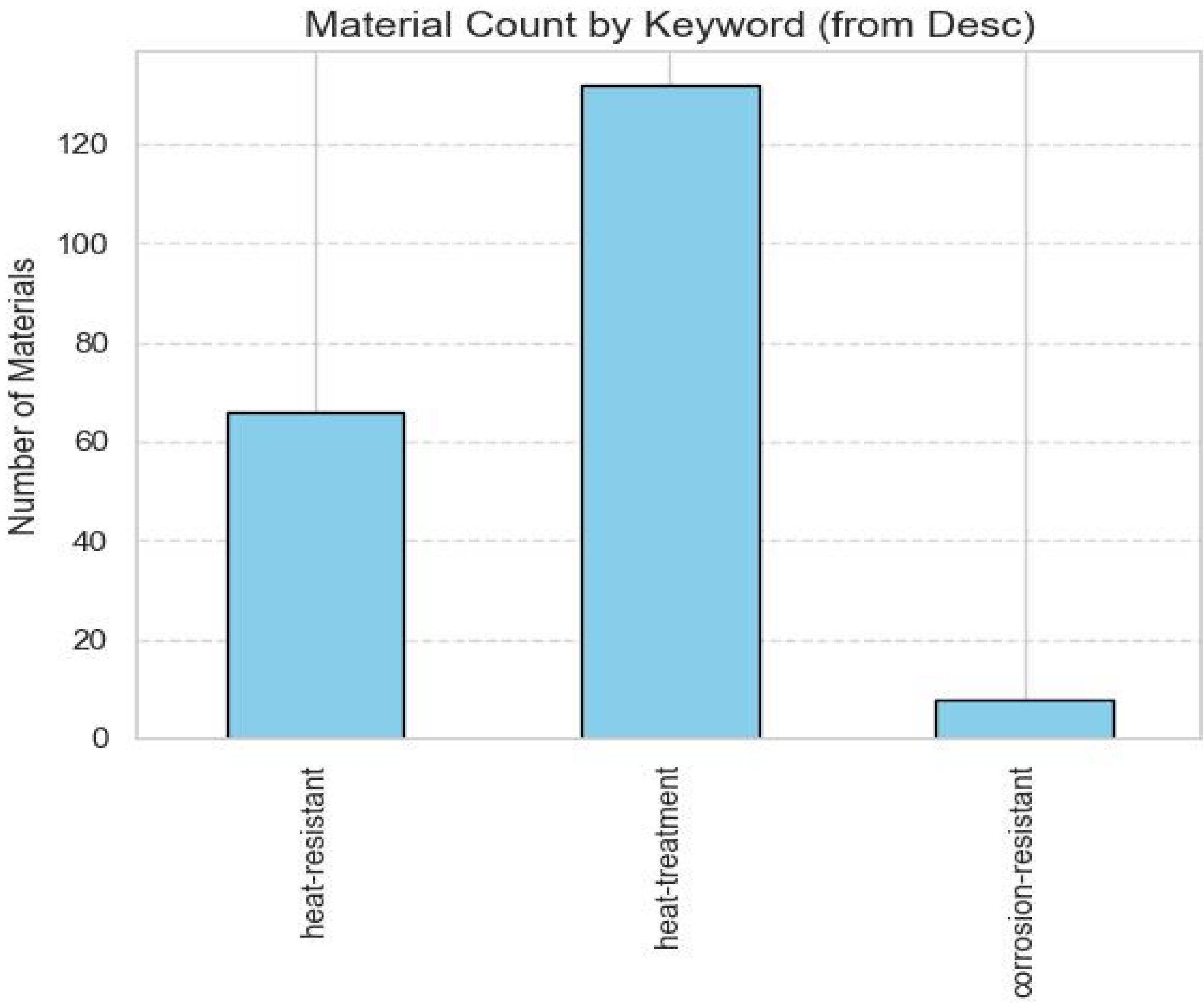
Steel 12Ch25N16G7AR: Moderate strength, highest ductility, top E and G.
Use Case: Marine/acid reactors, exhaust systems.

Variation by Heat Treatment:

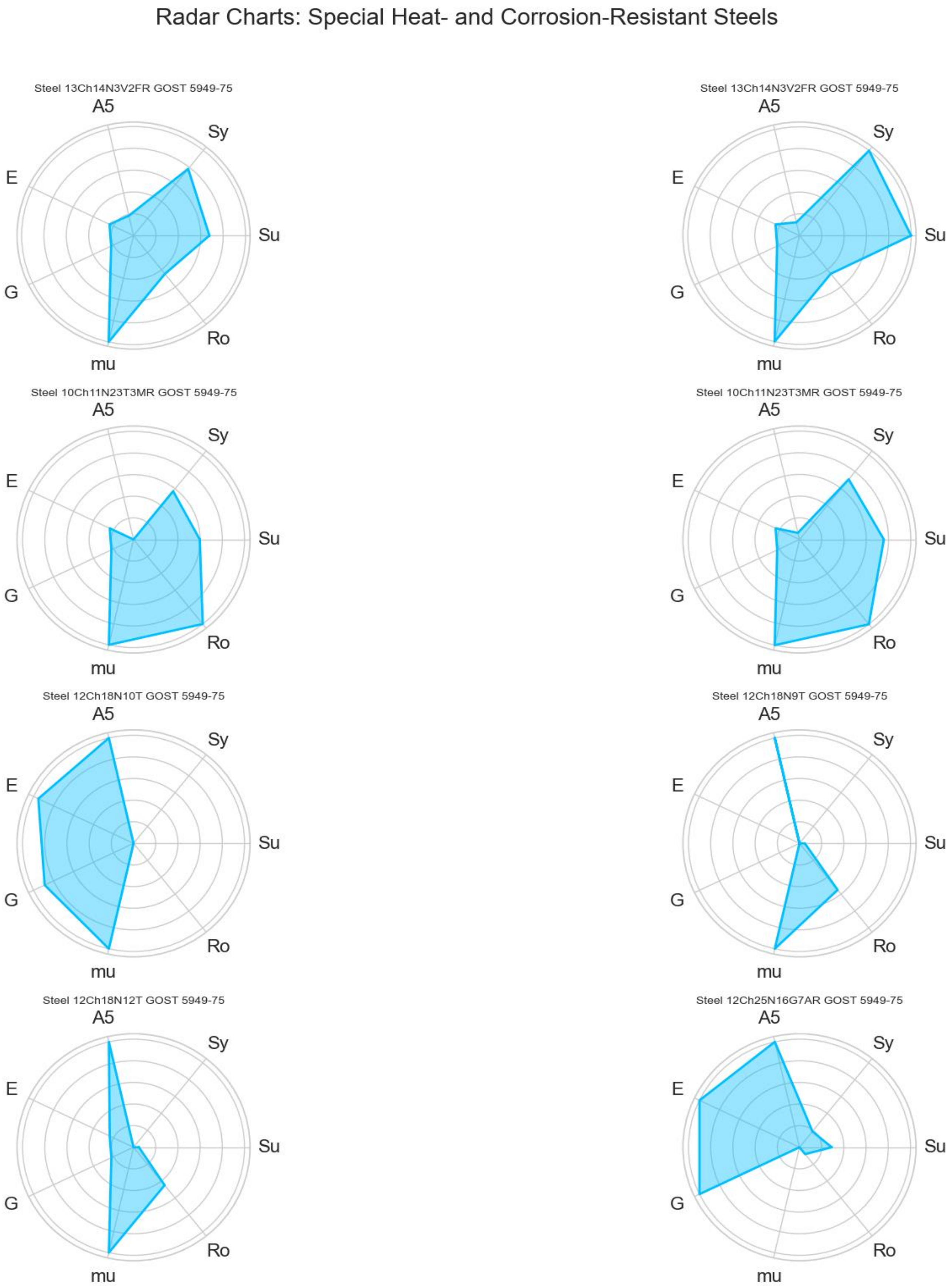
- **13Ch14N3V2FR:** Appears twice due to tempering difference (580°C vs 680°C). Stronger Sy in one variant.
- **10Ch11N23T3MR:** Also shown twice — high quench and temper (1100°C → 730–800°C) affects ductility/Su.

Summary Comparison Chart (8 Entries)

Steel	Strength	Ductility	Corrosion	Heat	Notes
13Ch14N3V2FR (1)	★★★★	★★★	★★★★	★★★★★	Top performance; gr temp loads
13Ch14N3V2FR (2)	★★★★	★★★	★★★★	★★★★★	Sy slightly low treatment effect vis
10Ch11N23T3MR (1)	★★★★	★★	★★★★	★★★★	Su+Sy stable; hea duty suitability
10Ch11N23T3MR (2)	★★★★	★★	★★★★	★★★★	Similar; more ducti sion 1
12Ch18N10T	★★	★★★★★	★★★★★	★★★★	Best ductility; grea resistance
12Ch18N9T	★★	★★★★	★★★★	★★★	Slightly stronger th
12Ch18N12T	★★	★★★★	★★★★	★★★★	Balanced; good cry performance
12Ch25N16G7AR	★★★	★★★★	★★★★★	★★★★	Best corrosion; wel tic



Final Engineering Takeaways



(Radar chart comparing all 8 materials across Su, Sy, A5, E, G, μ, Ro)

Heat Treatment Sensitivity:
Steels like 13Ch14N3V2FR and 10Ch11N23T3MR show property shifts depending on tempering. Enables tailored performance.

Elite Corrosion Resistance:
12Ch18N10T, 12Ch25N16G7AR, and 12Ch18N9T offer excellent corrosion resistance for marine/chemical/food use.

Formability–Strength Tradeoff:
As Su/Sy increases, A5 generally drops. Decide if flexibility or max load is your main concern.

Density & Cost:
Highly alloyed steels (e.g., 10Ch11N23T3MR) are heavier and pricier — best reserved for thermal/chemical extremes.