Security & SOC

Splunk Certified Cybersecurity Defense Analyst

KNOWLEDGE CHEATSHEET

MONITORING IMPROVES THROUGH FEEDBACK

Define Risk and goals → **Build** Collect telemetry → **Analyze** Test and report → **Review** Respond and refine

01 Security foundations and SOC operations

Protection goals and adversaries

- **Confidentiality** prevents unauthorized disclosure; **integrity** protects against unauthorized alteration or destruction; **availability** keeps information and services usable when needed.
- **Encryption and key rotation** protect confidentiality and limit exposure after compromise. **Checksums and cryptographic hashes** help detect changes; preserve a trusted reference value to validate integrity.
- **Personal data** can identify a person directly or indirectly. Names linked to addresses are a clear example; assess combinations of attributes and protect sensitive fields in collected logs.
- **Advanced persistent threats (APTs)** pursue sustained objectives with continued access. Motivations include espionage, financial gain, disruption and ideological causes; sophistication alone does not establish attribution.
- **Supply chain compromise** reaches victims through a trusted vendor, dependency or update. **Rootkits** conceal activity by manipulating operating-system behavior, including user-mode or kernel-mode interfaces.

Recognizing attack patterns

- **Social engineering** manipulates people. **Phishing** uses deceptive messages, **spearphishing** targets selected people, and **vishing** uses voice calls. Validate identity through an independent channel.
- **Credential stuffing** reuses breached username/password pairs. **Password spraying** tries a few passwords across many accounts; a **dictionary attack** tries likely password words or variants.
- **DoS** exhausts or disrupts a service; **DDoS** uses distributed sources. Repeated requests and unavailability warrant investigation, but one source IP or an HTTP success code alone cannot prove the cause.
- **Exfiltration** moves data out; **lateral movement** spreads access between internal systems; **C2** lets an adversary control compromised systems. Unusual outbound volume is a lead to validate against workload baselines.
- **Writable temporary directories** can stage malicious executables. Investigate parent process, command line, signer, hash and user context. Legitimate tools such as **rundll32** can proxy malicious execution.

Security & SOC / Threats & Hunting

Splunk Certified Cybersecurity Defense Analyst

KNOWLEDGE CHEATSHEET

MONITORING IMPROVES THROUGH FEEDBACK

Define Risk and goals → **Build** Collect telemetry → **Analyze** Test and report → **Review** Respond and refine

SOC responsibilities and measures

- **Analysts** triage and investigate; **security engineers** implement and maintain telemetry and detections; **architects** design solutions and choose tools; **SOC managers** direct people, priorities and performance.
- Escalate missing logs, broken field extractions and detection coverage gaps to the **engineering** function. Provide the affected source, expected evidence, time range and operational impact.
- **Blue teams** defend, **red teams** emulate adversaries, and **purple teams** coordinate offensive findings with defensive improvements. Use their results to refine detection coverage.
- **MTTD** measures detection delay; **MTTA** measures acknowledgment delay. **MTTR** can mean respond, resolve or recover: define the endpoint. For alert-to-disposition tracking, measure from alert creation to closure, including false positives.
- **Dwell time** describes how long an adversary remains present before the defined endpoint, often detection. Compare metrics using consistent clocks, severity groups and definitions; fast closure alone does not establish effective investigation.

Continuous monitoring cycle

- **Define and Predict** identifies risks, vulnerabilities and hypotheses. **Establish and Architect** sets monitoring requirements, frequency, architecture and responsibilities.
- **Implement and Collect** adds assets, identities, data sources, dashboards, searches and reports. Build useful drilldowns and verify that required telemetry is actually arriving.
- **Analyze and Report** evaluates collected evidence and search results against the hypothesis. Communicate findings, coverage gaps and opportunities to improve detection.
- **Respond and Review** mitigates identified problems and records lessons and recommendations, including after-action findings. Feed these back into the next cycle.

02 Threat frameworks and hunting

Frameworks and TTPs

- In **MITRE ATT&CK**, a **tactic** is the adversary's goal or why; a **technique** describes how the goal is achieved; a **procedure** is a specific implementation by an actor or tool.
- **Lateral movement** is a tactic; exploiting a remote service is a technique; using a named exploit against a particular SMB server is a procedure. **Brute force** belongs to Credential Access.

Threats & Hunting

Splunk Certified Cybersecurity Defense Analyst

KNOWLEDGE CHEATSHEET

A REPEATABLE THREAT HUNT

Hypothesis Testable activity → **Evidence** Relevant telemetry → **Analysis** Baselines and outliers → **Outcome** Findings and gaps

- The **Cyber Kill Chain** follows reconnaissance, weaponization, delivery, exploitation, installation, command and control, and actions on objectives. Boot persistence fits installation; stealing data serves an objective.
- **ATT&CK** organizes observed adversary behavior for intelligence, detection, emulation and assessment. **CIS Controls** prioritize safeguards; **NIST** publications guide security and risk practices; **ISO 27001** addresses information security management.
- **CMMC** assesses cybersecurity maturity for the defense industrial base. **PCI DSS** addresses payment-card security; **GDPR** addresses personal-data protection; **FISMA** concerns US federal information security. Match each framework to its purpose.

Intelligence and indicators

- **Strategic intelligence** describes business-level threats and trends for leadership. **Operational intelligence** concerns actor intent and campaigns; **tactical intelligence** supports defensive action through behaviors, techniques and indicators.
- **Indicators of compromise (IOCs)** include malicious hashes, domains, IPs and other observables with supporting context. Assess provenance, confidence, age and relevance before treating a match as malicious.
- The **Pyramid of Pain**, from easier to harder for attackers to change: hashes, IP addresses, domain names, network/host artifacts, tools, then TTPs. Behavior-based detections can impose more durable attacker costs.
- An IOC match is an **investigative lead**. Shared infrastructure, stale intelligence and legitimate administrative behavior can create misleading matches; corroborate with endpoint, identity and network evidence.

Hunt design and conclusions

- **Hypothesis-driven hunting** starts with a testable proposition about adversary behavior. **Data-driven hunting** starts by exploring available telemetry for suspicious patterns and developing explanations.
- Define the **scope**, expected artifacts, time window and required data before hunting. For suspected proxy execution and C2, correlate process telemetry, Sysmon or EDR, network flows and IDS observations.
- A hunt can succeed by testing a hypothesis, finding a visibility gap or improving a detection without finding an attacker. **No evidence found** is a bounded conclusion about the examined data, not proof of absence.
- A **baseline** models normal behavior for a relevant peer group and period. Account for seasonality, maintenance, job roles and changing workloads before classifying a deviation as suspicious.

Threats & Hunting / Data & Context

Splunk Certified Cybersecurity Defense Analyst

KNOWLEDGE CHEATSHEET

FROM RAW LOGS TO USEFUL CONTEXT

Collect Network and endpoint → **Normalize** CIM fields and tags → **Enrich** Assets and identities → **Apply** Security content

Hunt analysis methods

- **Least frequency of occurrence** and **long-tail analysis** investigate rare values, such as infrequent database users or unusual executables. Rare does not automatically mean malicious.
- **Anomaly and outlier detection** finds observations unlike peers or expected behavior. **Clustering** groups similar observations; dense email clusters can reveal campaigns with varying subjects, URLs or attachments.
- **Time-series analysis** examines changes, spikes and periodic behavior. **Co-occurrence analysis** examines items or events appearing together; **stack counting** groups values and counts their frequency.
- Turn useful hunt findings into **repeatable detections**, data-quality fixes or response steps. Document both evidence and limitations so another analyst can reproduce the conclusion.

03 Security data and context

Selecting the right telemetry

- **NetFlow and VPC Flow Logs** summarize communications such as addresses, ports, volume and accept/reject decisions. **PCAP** contains captured packets; available payload depends on capture settings and encryption.
- **Web proxy logs** show client web requests and destinations. **Web server access logs** show requests received by the server. **DNS logs** connect queried names with clients and answers.
- **IDS** detects and alerts; **IPS** can block inline. Firewalls and proxies add connection or policy context. Their visibility depends on placement, inspection capabilities and encrypted traffic handling.
- **EDR, Sysmon and endpoint logs** connect processes, files, registry activity and network connections. Authentication and IAM logs explain accounts, logins and privilege changes; email logs reveal delivery and phishing context.
- Validate **timestamps, timezone, retention, ingest delay and completeness** before correlating sources. A successful HTTP status or an absent alert is insufficient evidence of whether an attack succeeded.

Data & Context

Splunk Certified Cybersecurity Defense Analyst

KNOWLEDGE CHEATSHEET

FROM RAW LOGS TO USEFUL CONTEXT

Collect — Network and endpoint → **Normalize** — CIM fields and tags → **Enrich** — Assets and identities → **Apply** — Security content

CIM and data models

- The **Common Information Model (CIM)** standardizes field names, values and categorization across vendors so shared searches and data models can correlate equivalent events.
- Use appropriate **Splunk add-ons** for parsing and search-time knowledge. Validate field extractions, aliases, calculated fields, event types and **tags**; ingesting raw logs alone does not make them CIM compliant.
- A **data model** organizes constrained datasets and fields for a class of activity. Events must satisfy dataset constraints: an authentication event with good fields can still be excluded if its required authentication tag is missing.
- Use **Authentication** for login and privilege activity, **Network Traffic** for flows, **Web** for web transactions and **Endpoint** for process, filesystem and registry activity. Endpoint filesystem fields include **file_acl**.
- **src/src_ip** identifies a source and **dest/dest_ip** a destination when provided by the dataset. **src_user** identifies the initiating user in applicable authentication events. Splunk **host** metadata identifies the event's host, which can differ from either network endpoint.

Enrichment and content discovery

- The **Asset and Identity framework** enriches events with known device and user context, including ownership, priority and categories. Asset matching can add fields such as **src_category** and **dest_category**.
- Maintain asset IP, hostname and identity mappings to connect network evidence to real entities. Use a **lookup** backed by suitable reference data to add a public IP's **ASN** and owner; field extraction cannot derive ownership information that is absent.
- **Splunk Security Essentials (SSE)** helps inventory data and discover, evaluate and deploy security content. Run **Data Inventory** introspection or configure inventory manually before filtering content by available data.
- In SSE, the **Data Availability filter** narrows use cases by required data; the **Data Availability dashboard** monitors ingest latency. **Content Mapping** tracks deployed content, and **Security Data Journey** guides maturity and onboarding.
- **Enterprise Security Content Update (ESCU)** delivers security analytics and Analytic Stories for related use cases. **Splunk Answers** provides community answers; **Splunk Lantern** supplies use-case guidance; product documentation defines command and configuration behavior.

SPL Analysis

Splunk Certified Cybersecurity Defense Analyst

KNOWLEDGE CHEATSHEET

A FOCUSED INVESTIGATION SEARCH

Scope Time and source → **Shape** Extract and enrich → **Summarize** Counts and patterns → **Inspect** Outliers and evidence

04 SPL for investigation

Search scope and performance

- Set a relevant **time range** and filter early with **index**, **sourcetype**, host and selective terms. Avoid unnecessary All Time searches and leading wildcards on large datasets.
- Run selective **distributable streaming commands** before centralized work where semantics allow, so indexers can reduce data before it reaches the search head. Command order must preserve the intended result.
- Use **fields** to retain needed fields during analysis and **table** for final presentation. After **stats**, only aggregation outputs and grouping fields remain; include host or user in the grouping if later steps need them.
- Prefer **stats** for grouped counts, sums and time ranges when sufficient. **transaction** groups related events and retains richer event context, but can be more expensive; choose it only when those semantics are needed.
- **Fast** mode limits field discovery; **Smart** adapts to search type; **Verbose** discovers more fields and retains event detail for investigation. Missing displayed fields may reflect search mode, extraction or permissions.
- Use **TERM(value)** to match an indexed term containing minor breakers, subject to tokenization rules. **CASE(value)** requests case-sensitive matching. Inspect available sourcetypes with `| metadata type=sourcetypes`.

Fields and transformations

- **rex** extracts fields using regular expressions: `| rex field=_raw "user=(?<account>[^ ]+)"`. **regex** filters events using a regular expression; it does not create an extracted field.
- **eval** creates or changes fields using expressions, including text, JSON, comparison and conditional functions. `| eval src=coalesce(src,machine_name)` chooses the first non-null value; an empty string can require separate normalization.
- Use **rex mode=sed** for substitutions, such as `| rex field=account mode=sed "s/[0-9]/X/g"`. Search-time masking changes the returned field, not the stored indexed event.
- **foreach** iterates an expression over fields matching a wildcard list, for example `| foreach bytes_* [ eval <<FIELD>>=coalesce('<<FIELD>>',0) ]`. Its bracketed template operates on fields in the current results.
- Start a small synthetic example with **makeresults**, then add values using **eval**. Use **lookup** for external reference enrichment and **rename** for readable output names.

SPL Analysis

Splunk Certified Cybersecurity Defense Analyst

KNOWLEDGE CHEATSHEET

A FOCUSED INVESTIGATION SEARCH

Scope Time and source → **Shape** Extract and enrich → **Summarize** Counts and patterns → **Inspect** Outliers and evidence

Aggregation and outliers

- **stats count by user** produces one row per user. To examine high counts: `| stats count by user | where count>1000 | sort - count`. **where** evaluates expressions over the current result fields.
- For rejected flows, filter the action first, then aggregate: `| stats count as blocked dc(dest_ip) as destinations values(dest_ip) as dest_ips by src_ip | where destinations>5 AND blocked>1000 | table src_ip dest_ips`.
- **dc(field)** counts distinct values; **values(field)** returns distinct values as a multivalue field. Distinct IP addresses measure addresses, which may not map one-to-one to people.
- **top** returns common values and **rare** returns uncommon values. Compare counts within an appropriate peer group, time window and denominator before identifying an outlier.
- **bin _time span=1d** groups timestamps into daily buckets; **timechart** summarizes over time. **streamstats** computes running aggregates in result order, so sort chronologically before calculating cumulative values.
- For cumulative distinct clients, retain each day's client values and use **streamstats dc(ips)** in chronological order. Summing daily distinct counts overcounts clients that reappear on several days.

Indexed and accelerated searches

- **tstats** calculates statistics from indexed fields in **tsidx** structures or from accelerated data-model summaries. It avoids scanning raw events for each calculation; arbitrary search-time fields are not automatically available.
- **Data model acceleration** builds summaries for eligible datasets. CIM normalization supplies consistent meaning; acceleration supplies a faster access path. Validate both dataset membership and summary coverage.
- **summariesonly=true** restricts tstats to acceleration summaries and can omit data outside their coverage or not yet summarized. Include a deliberate time window and account for summary lag.
- To summarize risk by entity: `| tstats summariesonly=true sum(All_Risk.calculated_risk_score) as total_risk from datamodel=Risk.All_Risk by All_Risk.risk_object`. Verify the accelerated model and fields are available.

Investigation & Response

Splunk Certified Cybersecurity Defense Analyst

KNOWLEDGE CHEATSHEET

FROM OBSERVATION TO RESPONSE

Detect Correlation and risk → **Triage** Context and owner → **Investigate** Timeline and pivots → **Respond** Actions and playbooks

05 Enterprise Security investigation and response

Detections and framework context

- A **correlation search** evaluates security conditions and can trigger actions. The **Notable** adaptive response action creates a notable for analyst triage; a search result becomes an alert only through the configured detection workflow.
- The **Notable Event framework** supports ownership, urgency, status and triage in **Incident Review**. Investigation workflows collect related findings and evidence into a larger incident narrative.
- **Annotations** attach framework context such as MITRE ATT&CK techniques, Kill Chain stages, CIS and NIST mappings to detection results. These mappings explain analytic intent and coverage; they do not prove compliance or attacker attribution.
- The **Threat Intelligence framework** manages intelligence for matching observables. **Asset and Identity** supplies entity context; the **Risk framework** accumulates suspicious behavior; **Adaptive Response** integrates actions inside or outside Splunk.

Risk-based alerting

- **Risk-based alerting (RBA)** combines observations across time and detections rather than requiring every weak signal to produce a standalone incident. Risk incident rules identify patterns or thresholds warranting investigation.
- A **risk object** is the entity being assessed, such as a user or system. **risk_object_type** describes its kind; **risk_score** quantifies an observation's assigned risk. Risk scores are prioritization signals, not probabilities of compromise.
- **Risk factors** adjust scores according to conditions such as asset priority or identity context. Consistent entity naming and normalization prevent one user or device from fragmenting across several risk objects.
- A useful **risk_message** explains the activity using relevant event fields. Include the context needed for triage; extracting a field makes it searchable, while adding it to the message makes the observation easier to interpret.
- Open the **Risk Event Timeline** from a risk notable's risk-event count in Incident Review. Review chronology, contributing detections, scores and **threat objects** such as domains, IPs or command lines; follow drilldowns to original evidence.

Investigation & Response

Splunk Certified Cybersecurity Defense Analyst

KNOWLEDGE CHEATSHEET

FROM OBSERVATION TO RESPONSE

Detect Correlation and risk → **Triage** Context and owner → **Investigate** Timeline and pivots → **Respond** Actions and playbooks

Triage, dispositions and dashboards

- **True positive**: malicious activity correctly detected. **False positive**: a detection incorrectly indicates its target condition. **False negative**: relevant activity was missed. **True negative**: the target condition was absent and no alert fired.
- A **benign positive** identifies activity that occurred but is expected or authorized. Record why it is benign and consider narrowly scoped tuning. Missing investigation evidence does not establish either maliciousness or innocence; document the unresolved outcome.
- Triage by **urgency, affected asset, identity, evidence and scope**. Assign an owner, preserve a timeline, record searches and justify disposition. Validate containment and recovery rather than closing solely because an action was launched.
- **Access** dashboards cover authentication, login and account activity; **Access Anomalies** supports unusual access investigations such as impossible travel. Check VPNs, geolocation quality and shared accounts before concluding compromise.
- **Identity Center** supports identity and watchlist context; **New Domain Analysis** helps investigate newly observed or suspicious domains, including typosquatting. **Endpoint** dashboards include host and vulnerability context; select views according to the evidence needed.

Response actions and automation

- **Workflow actions** pivot from fields or events into searches or external tools, such as an IP reputation query. A **drilldown** opens the evidence behind a dashboard result or notable.
- **Adaptive Response actions** run configured actions from detections or supported investigation workflows. Engineers can expose appropriate actions, such as notification email or an integrated response, for analysts to invoke.
- A **SOAR playbook** automates a sequence such as enrich an indicator, evaluate the result, request or perform containment, and update the original finding. Trigger it through the configured detection or Adaptive Response integration.
- Use **runbooks** to document human procedures and **playbooks** to execute supported automation. Validate permissions, approval steps, action results and audit records when automating disruptive containment.