Massimo Ivaldi

A SEARCH COMPANY

GUIDA PRATICA

Le 10 regole fondamentali della ricerca interna

Guida pratica, valida per qualunque motore, per configurare una ricerca che fa davvero trovare



massimoivaldi.com

AUTORE Massimo Ivaldi

Le 10 regole fondamentali della ricerca interna

Sintesi

Una ricerca interna funziona quando l'utente trova ciò che cerca al primo tentativo, anche scrivendo male, con parole sue e senza sapere come sono organizzati i contenuti. Il motore conta meno della configurazione: quasi tutti gli strumenti sul mercato offrono le stesse leve di base, e la differenza la fa come vengono regolate.

Questo documento raccoglie dieci regole valide per qualunque motore:

1. Decidere cosa è ricercabile, cosa è filtrabile e cosa si mostra.
2. Tollerare errori di battitura e varianti linguistiche.
3. Trattare la pertinenza come una decisione di business, non come un default.
4. Offrire filtri, faccette e ordinamenti coerenti con il modo in cui l'utente sceglie.
5. Rispondere subito e mostrare perché un risultato è pertinente.
6. Non lasciare mai l'utente davanti a una pagina vuota.
7. Restituire risultati puliti: senza doppioni, raggruppati e paginati.
8. Mostrare a ciascuno solo ciò che può vedere.
9. Misurare cosa cercano gli utenti e correggere di conseguenza.
10. Tenere l'indice aggiornato e il sistema stabile.

Introduzione: regole, non funzionalità

Chi usa la casella di ricerca ha un'intenzione precisa e poca pazienza. Su un sito di e-commerce, un portale documentale o una intranet, la ricerca è spesso il percorso più breve verso una vendita, una risposta o un documento. Quando fallisce, l'utente non pensa "il contenuto non c'è": pensa "qui non si trova niente", e se ne va.

Le schede tecniche dei motori di ricerca elencano decine di funzionalità con nomi diversi. Sotto i nomi, però, i concetti sono quasi sempre gli stessi: un indice, regole per confrontare la query con i documenti, regole per ordinarli, strumenti per restringerli. Questo documento separa ciò che è specifico di un prodotto da ciò che è un principio generale, e si concentra sul secondo.

L'obiettivo è pratico: dare a chi configura una ricerca interna, di qualunque tipo, una lista di decisioni da prendere consapevolmente invece di lasciarle ai valori predefiniti.

Regola 1 — Decidere il ruolo di ogni campo

Ogni campo di un documento ha un ruolo diverso, e va dichiarato esplicitamente. Indicizzare tutto "per sicurezza" peggiora la pertinenza: una parola trovata in un codice interno o in una nota tecnica fa salire risultati che l'utente non voleva.

Per ciascun campo vanno decise quattro cose:

Ruolo	Domanda da porsi	Esempio
Ricercabile	L'utente scriverebbe questo testo nella casella?	Titolo, descrizione, marca
Filtrabile	L'utente vuole restringere i risultati su questo valore?	Categoria, prezzo, data, lingua
Ordinabile	L'utente vuole ordinare su questo valore?	Prezzo, data, popolarità
Visualizzato	Serve mostrarlo nella pagina dei risultati?	Titolo, immagine, estratto

Tre accorgimenti valgono quasi sempre:

- **Pochi campi ricercabili, in ordine di importanza.** Il titolo pesa più della descrizione, la descrizione più del corpo del testo.
- **Niente dati sensibili o interni nell'indice** se non servono alla ricerca: ciò che è indicizzato può emergere in un risultato.
- **Dati preparati prima dell'indicizzazione.** Testo ripulito da markup, valori normalizzati (stesso formato per date, unità, marche), documenti lunghi spezzati in sezioni se l'utente cerca risposte puntuali.

Regola 2 — Capire l'utente anche quando scrive male

Gli utenti sbagliano a scrivere, usano sinonimi, abbreviazioni, singolari e plurali. Una ricerca che pretende la parola esatta fallisce proprio sulle query più comuni.

Le leve di base, presenti in quasi tutti i motori:

- **Tolleranza agli errori di battitura.** Accetta parole che differiscono di una o due lettere. Va calibrata: sulle parole brevi un errore cambia il significato, e sui codici prodotto o sui numeri va spesso disattivata.
- **Normalizzazione.** Maiuscole, accenti e punteggiatura non devono contare: "perche", "perché" e "PERCHÉ" sono la stessa ricerca.

- **Varianti morfologiche.** Riportare le parole alla radice ("scarpa" e "scarpe", "configurare" e "configurazione"). Dipende dalla lingua: va impostata quella giusta per ogni contenuto.
- **Sinonimi.** Un dizionario curato a mano, alimentato dalle ricerche reali degli utenti: "cellulare" e "smartphone", "ferie" e "congedo", sigle e loro espansioni.
- **Parole vuote.** Articoli e preposizioni ("il", "di", "per") pesano poco o nulla nel confronto, ma vanno trattati con cautela dove contano ("essere o non essere").
- **Ricerca per prefisso.** Restituire risultati mentre l'utente sta ancora scrivendo la parola.
- **Frase esatta.** Le virgolette permettono all'utente esperto di cercare una sequenza precisa di parole.

La regola pratica: tolleranza ampia sul testo libero, rigore su codici, numeri e identificativi.

Regola 3 — La pertinenza è una scelta, non un default

L'ordine dei risultati è la decisione più importante della configurazione, perché l'utente guarda quasi solo i primi. Ogni motore ha un ordinamento predefinito basato sul testo; nessun motore conosce le priorità della vostra organizzazione.

I criteri testuali di base sono simili ovunque:

- quante parole della query compaiono nel documento;
- quanti errori di battitura sono serviti per trovarle;
- quanto sono vicine tra loro;
- in quale campo compaiono (titolo prima del corpo);
- se la corrispondenza è esatta o parziale.

A questi si aggiungono i **segnali di business**, che vanno definiti da chi conosce il dominio: popolarità, disponibilità, margine, data di pubblicazione, autorevolezza della fonte. Alcuni motori li combinano in un punteggio, altri in una sequenza ordinata di criteri; il principio è lo stesso.

Due strumenti complementari:

- **Risultati fissati.** Per poche query strategiche, imporre un risultato in cima ("orari" → pagina degli orari).
- **Penalizzazioni.** Abbassare contenuti obsoleti, esauriti o archiviati senza eliminarli.

Ogni modifica va verificata su un insieme fisso di query di prova prima di andare in produzione: un ritocco che migliora una ricerca spesso ne peggiora altre.

Regola 4 — Aiutare a restringere, non solo a cercare

Molti utenti partono da una parola generica ("giacca", "contratto") e poi restringono. Filtri e faccette trasformano una lista lunga in una scelta gestibile.

- **Filtri.** Condizioni su valori strutturati: categoria, intervallo di prezzo, data, autore, lingua. Devono funzionare insieme alla ricerca testuale, non al suo posto.
- **Faccette.** I valori disponibili per ogni filtro, con il numero di risultati accanto ("Rosso (12)"). Mostrano all'utente cosa c'è prima che clicchi, ed evitano combinazioni che danno zero risultati.
- **Ordinamento.** Per pertinenza come default; per prezzo, data o popolarità quando l'utente lo chiede. Ordinare per data distrugge la pertinenza: va offerto, non imposto.
- **Filtri geografici.** Dove la posizione conta (negozi, sedi, servizi), quasi tutti i motori permettono di filtrare o ordinare per distanza.

Buone pratiche comuni:

- poche faccette, scelte in base a come gli utenti decidono davvero;
- valori puliti e coerenti (una marca scritta in tre modi diventa tre voci);
- filtri attivi sempre visibili e facili da rimuovere.

Regola 5 — Rispondere subito e mostrare il perché

La velocità è parte della pertinenza: una risposta lenta viene percepita come una risposta sbagliata. Le ricerche "mentre scrivi" richiedono tempi di risposta nell'ordine delle decine di millisecondi per sembrare istantanee.

Per la pagina dei risultati valgono alcune regole comuni:

- **Evidenziazione.** Le parole trovate vanno messe in risalto: l'utente capisce a colpo d'occhio perché un risultato è lì.
- **Estratto.** Per documenti lunghi, mostrare il passaggio che contiene i termini cercati, non le prime righe.
- **Solo i campi utili.** Restituire al client solo ciò che si mostra: risposte più leggere, più veloci, più sicure.
- **Suggerimenti in tempo reale.** Completamento della query e anteprima dei primi risultati mentre si digita.
- **Coerenza tra dispositivi.** La stessa ricerca deve dare gli stessi risultati su desktop e mobile; cambia solo la presentazione.

Regola 6 — Mai un vicolo cieco

Una pagina "nessun risultato" è il punto in cui la ricerca perde più utenti. Va trattata come un caso da progettare, non come un errore.

Le strategie, dalla più semplice alla più elaborata:

1. **Rilassare la query.** Se nessun documento contiene tutte le parole, mostrare quelli che ne contengono la maggior parte, segnalandolo all'utente. Molti motori lo fanno in automatico, togliendo prima le parole meno importanti.
2. **Proporre una correzione.** "Forse cercavi…" quando la query sembra un errore di battitura.
3. **Suggerire alternative.** Ricerche popolari, categorie principali, contenuti correlati.
4. **Offrire un'uscita.** Un contatto, un modulo, un link alla navigazione per categorie.

Ogni ricerca senza risultati va registrata: è la lista più preziosa per sapere quali sinonimi aggiungere e quali contenuti mancano.

Regola 7 — Risultati puliti e navigabili

Una lista piena di doppioni o di varianti dello stesso oggetto fa sembrare la ricerca povera anche quando è precisa.

- **Deduplicazione.** Lo stesso prodotto in dieci taglie, o lo stesso documento in tre versioni, va mostrato una volta sola. Quasi tutti i motori permettono di raggruppare i risultati per un campo (codice articolo, documento padre) e restituirne uno per gruppo.
- **Varianti dentro il risultato.** Taglie, colori o versioni si mostrano nella scheda, non come risultati separati.
- **Paginazione o scorrimento.** Il numero totale di risultati va comunicato, anche se approssimato. Oltre qualche centinaio di risultati nessuno scorre: è meglio spingere verso i filtri.
- **Ricerca su più fonti.** Quando i contenuti stanno in indici diversi (prodotti, articoli, FAQ), si possono interrogare insieme e mostrare in sezioni separate, oppure fondere in un'unica lista. Alcuni motori lo fanno nativamente, altri richiedono di farlo nell'applicazione.

Regola 8 — Ognuno vede solo ciò che può vedere

La ricerca è un modo rapidissimo per far trapelare informazioni: un titolo o un estratto in un risultato basta a rivelare un documento riservato. In una intranet o in un servizio con più clienti, i permessi vanno applicati dentro la ricerca, non dopo.

- **Filtri di sicurezza lato server.** Ogni query porta con sé un filtro (reparto, cliente, ruolo) che l'utente non può modificare. Filtrare i risultati nell'interfaccia, dopo averli ricevuti, non basta: anche i conteggi e le faccette rivelerebbero ciò che non si vede.
- **Chiavi con privilegi minimi.** Chiavi di sola lettura per il pubblico, chiavi di scrittura solo per i sistemi che indicizzano, chiavi amministrative mai esposte nel browser.
- **Separazione dei clienti.** In un servizio condiviso, indici separati o un filtro obbligatorio per cliente. Molti motori offrono chiavi o token che incorporano il filtro.
- **Permessi aggiornati.** Quando un utente perde un accesso, la ricerca deve riflettere il cambiamento in tempi brevi.

Regola 9 — Misurare, correggere, ripetere

Una ricerca non si configura una volta per tutte: si migliora leggendo cosa fanno gli utenti. Le query sono la voce del cliente, scritta con le sue parole.

Indicatore	Cosa misura	Cosa fare se peggiora
Ricerche senza risultati	Query che non trovano nulla	Aggiungere sinonimi, contenuti mancanti, correzioni
Tasso di clic	Quante ricerche portano a un clic	Rivedere ordinamento e presentazione
Posizione media del clic	Quanto in basso si trova ciò che serve	Rivedere i criteri di pertinenza
Ricerche riformulate	Quante volte l'utente riprova subito	Studiare la prima query: cosa mancava?
Abbandono dopo la ricerca	Uscite senza clic	Controllare risultati vuoti o irrilevanti
Conversione da ricerca	Acquisti, download, ticket evitati	Misura l'impatto di business
Tempo di risposta	Latenza percepita	Ridurre campi, ottimizzare indice e infrastruttura

Un metodo semplice e diffuso:

1. Ogni mese, estrarre le query più frequenti e quelle senza risultati.
2. Per le prime 50, controllare a mano se i primi risultati sono quelli giusti.
3. Correggere sinonimi, pesi e risultati fissati.

4. Rifare il controllo sullo stesso insieme di query per verificare di non aver peggiorato nulla.

Regola 10 — Un indice fresco e un sistema stabile

Un risultato che porta a un prodotto esaurito o a una pagina cancellata costa fiducia quanto un risultato sbagliato.

- **Aggiornamenti incrementali.** Indicizzare solo ciò che cambia, appena cambia, invece di ricostruire tutto ogni notte. Prezzi e disponibilità richiedono tempi più stretti dei testi.
- **Cancellazioni propagate.** Ciò che sparisce dalla fonte deve sparire dall'indice.
- **Ricostruzioni senza interruzioni.** Quando serve reindicizzare tutto (nuove impostazioni, nuovo schema), costruire un indice nuovo accanto al vecchio e scambiarli alla fine.
- **Impostazioni versionate.** Sinonimi, pesi e campi vanno salvati come configurazione, non modificati a mano in produzione: così si può tornare indietro.
- **Ambiente di prova.** Ogni modifica di pertinenza si prova su una copia prima di applicarla.
- **Copie e backup.** Repliche per reggere i picchi e un backup regolare dell'indice e delle impostazioni.
- **Monitoraggio.** Latenza, errori, code di indicizzazione in ritardo: allarmi prima che se ne accorgano gli utenti.

Le funzioni di base, a colpo d'occhio

Quasi tutti i motori di ricerca offrono queste funzioni, a volte con nomi diversi. La tabella le collega alla regola che servono.

Funzione	A cosa serve	Regola
Campi ricercabili e loro peso	Cercare solo dove ha senso, nell'ordine giusto	1, 3
Campi visualizzati	Restituire solo ciò che si mostra	1, 5
Tolleranza agli errori	Trovare anche con errori di battitura	2
Analisi linguistica (radici, accenti)	Trattare varianti della stessa parola come una	2
Sinonimi	Collegare parole diverse con lo stesso significato	2, 6
Parole vuote	Ignorare articoli e preposizioni	2

Funzione	A cosa serve	Regola
Ricerca per prefisso	Risultati mentre si scrive	2, 5
Frase esatta	Cercare una sequenza precisa	2
Regole di ordinamento	Decidere chi viene prima	3
Risultati fissati	Imporre la risposta a query chiave	3
Filtri	Restringere per valori strutturati	4, 8
Faccette con conteggi	Mostrare le scelte disponibili	4
Ordinamento per campo	Prezzo, data, popolarità	4
Ricerca per distanza	Risultati vicini a un punto	4
Evidenziazione ed estratti	Mostrare perché un risultato è pertinente	5
Raggruppamento per campo	Eliminare doppioni e varianti	7
Ricerca su più indici	Interrogare fonti diverse insieme	7
Chiavi e token con permessi	Limitare chi vede cosa	8
Statistiche di ricerca	Capire cosa cercano gli utenti	9
Indicizzazione incrementale	Tenere l'indice aggiornato	10

Alcuni motori aggiungono funzioni più avanzate, come la **ricerca semantica** (che confronta il significato invece delle parole) o la **ricerca ibrida** (che combina significato e parole chiave). Sono utili quando gli utenti descrivono un bisogno invece di nominare un oggetto, ma non sostituiscono le regole precedenti: si costruiscono sopra.

Checklist e conclusioni

Prima di mettere in produzione una ricerca interna, ogni voce dovrebbe avere una risposta consapevole:

- [ ] Per ogni campo è deciso se è ricercabile, filtrabile, ordinabile, visualizzato
- [ ] I campi ricercabili sono in ordine di importanza
- [ ] La lingua dei contenuti è impostata e la tolleranza agli errori è calibrata (disattivata su codici e numeri)
- [ ] Esiste un primo dizionario di sinonimi
- [ ] I criteri di ordinamento includono i segnali di business
- [ ] Le query strategiche hanno un risultato fissato

- [ ] Filtri e faccette riflettono come gli utenti scelgono, con valori puliti
- [ ] Le parole trovate sono evidenziate e gli estratti sono pertinenti
- [ ] La pagina senza risultati propone alternative
- [ ] I doppioni sono raggruppati
- [ ] I permessi sono applicati lato server, con chiavi a privilegi minimi
- [ ] Le ricerche e i clic vengono registrati
- [ ] C'è un insieme fisso di query di prova
- [ ] L'indice si aggiorna in modo incrementale e le impostazioni sono versionate

Il motore scelto conta meno di quanto sembri. Quasi tutti offrono le stesse leve; la qualità della ricerca dipende da chi le regola, dalla pulizia dei dati e dall'abitudine di guardare ogni mese cosa cercano davvero gli utenti.