

If you are working with the Wikidata + Google Knowledge Graph MCP server, `kg_related` is the tool you reach for when a name match is not enough and a single entity record is not enough either. It sits in the space between search and inspection. `kg_search` helps you find candidate entities. `kg_entity` helps you read selected facts about one entity. `kg_related` becomes useful when the hard part is context, not retrieval.

That distinction matters more than it sounds. In real entity work, the biggest source of bad matches is not missing data. It is missing neighborhood. A label can look right, an occupation can look right, even a birth year can look right, and yet the record is still wrong because it lives in the wrong cluster of people, places, works, or organizations. A politician belongs to one party, serves in one legislature, and is tied to a specific region. A musician belongs to a scene, collaborates with certain artists, and releases work through a pattern of labels or ensembles. A company appears in an industry web with founders, subsidiaries, and headquarters. Context settles arguments that labels never can.

That is when `kg_related` earns its keep.

What `kg_related` is really for

The project documentation names `kg_related` as one of the MCP tools exposed by the server, alongside `kg_search`, `kg_entity`, `kg_resolve`, and `kg_status`. The overall server is built to let AI agents search Wikidata, read selected facts, and link local records to Wikidata QIDs with inspectable evidence and explicit uncertainty when evidence is insufficient. That design philosophy tells you how **Wikidata MCP dataset** to think about `kg_related`.

It is not a broad crawler and not a license to pull in every neighboring node. The server emphasizes bounded search across the product. By default it returns three candidates and caps at five rather than dumping large raw result sets. That restraint is important. It means relatedness is intended to support judgment, not overwhelm it.

In practice, `kg_related` is the tool for answering questions like these in a controlled way:

Does this candidate sit near the people or institutions I would expect?

Is this entity related to the same place, field, period, or organization as my local record suggests?

If I already have a likely QID, what surrounding entities would help me verify it without manually traversing many hops?

Those are different from pure lookup questions. If you already know the exact QID and only need a birth date, an official website, or a handful of selected statements with qualifiers and references, `kg_entity` is cleaner. If you do not know the entity yet, `kg_search` or `kg_resolve` comes first. `kg_related` is for the middle stretch, when you need confirming structure.

The pattern I see most often in real workflows

A surprisingly common mistake in MCP for Google Knowledge Graph and Wikidata workflows is trying to resolve identity in one move. People search a label, skim the top candidate, and accept it because the title looks plausible. That works for distinctive names. It fails on common names, translated names, stage names, institutions with nearly identical names, and records with sparse local metadata.

The safer pattern has three beats.

First, search for plausible candidates. Second, inspect one or two high-signal facts on each candidate. Third, use related entities to see whether the candidate belongs in the right neighborhood.

That third step is where `kg_related` adds value. It is the difference between “this might be Jane Smith” and “this is the Jane Smith who served in the same legislature, belongs to the same party, and is tied to the same district that my source record mentions.”

The benefit is not just better matching. It also makes your evidence inspectable. This project is built around explicit uncertainty and deterministic outcomes such as `AUTO_MATCH`, `HOLD`, `AMBIGUOUS`, and `NO_CANDIDATE`. Related entities often provide the extra evidence needed to move a record out of `HOLD` or `AMBIGUOUS`, but only when the surrounding graph actually lines up. If it does not, that negative signal is just as useful.

When `kg_related` is the right next step

You should think about `kg_related` as a context amplifier. It is most useful when you have a candidate, or a short list of candidates, and need to test whether one of them coheres with your local record.

Here are the situations where it tends to pay off:

- the label is common or shared by several notable entities
- your local record has relational clues, such as employer, genre, party, collaborator, or location
- you need inspectable evidence for a match decision rather than a guess based on one field
- `kg_resolve` returns uncertainty and you want to understand why a candidate is close but not safe
- you are preparing a human review queue and want richer context without opening a full manual research session

That short list captures most practical cases. The common thread is that the identity question cannot be settled by surface text alone.

Consider a local person record with the name “Michael Brown,” a mention of a state senate, and a rough time period. A basic search might produce multiple QIDs. A direct entity read can tell you each candidate’s office or party, but related entities might expose which candidate is connected to the correct state, legislature, or colleagues. One candidate may sit in an entirely different political network. That is often enough to discard it with confidence.

The same principle applies to organizations. A local record might say “St. Mary’s College” with almost no other metadata. Search results may produce schools in different countries or denominational traditions. A single fact like instance of or inception year may not decide it. Related entities, however, might reveal the city, affiliated institutions, or associated people that place the correct candidate in the right local ecosystem.

Why relatedness matters more than extra facts

A lot of users reach for more properties when they are uncertain. Sometimes that helps. The server supports selected-fact retrieval, including ranks, qualifiers, and references on request, which is valuable when you need to inspect a statement carefully. But there is a limit to how much confidence you can squeeze from one entity in isolation.

Relatedness does something different. It reveals whether an entity behaves like the one you think it is.

That difference is subtle but important. A person record may have the right occupation and the right birth year, yet still be wrong because the local record refers to a different national context. A place may share a name with another place, and both may have similar administrative types. The neighboring entities tell you whether the broader identity fits.

In MCP for wikidata work, that is often the point where machine assistance starts to resemble expert research. Human researchers rarely stop at labels. They ask, “Who else is connected here?” `kg_related` gives an MCP client a structured way to do the same without pretending that graph proximity is proof.

The Google Knowledge Graph angle, and why it should stay secondary

The project supports an optional Google cross-check through exact ID joins, using `/m/` for Wikidata property P646 and `/g/` for P2671. That is useful, but the documentation is careful about how to interpret it. Agreement between Google and Wikidata is provider concordance, not proof of identity.

That caution should shape how you use `kg_related` in MCP for google knowledge graph and wikidata settings.

If you already have a strong Wikidata candidate and its surrounding relationships fit your local record, a Google cross-check can strengthen your comfort that multiple providers are talking about the same thing. It is especially handy when your downstream system already stores Google-style IDs or when a workflow wants to compare provider coverage.

What it should not do is replace relational judgment. If the candidate’s network looks wrong, provider agreement does not rescue the match. Two systems can concord on the same mistaken interpretation for your use case, or they can both point to a broadly named concept that is too coarse for your local record. The server’s own documentation avoids treating provider agreement as identity proof, and that is the right stance.

This is one reason `kg_related` remains relevant even when the optional Google API is available. The value of relatedness is not redundancy between providers. The value is contextual fit.

How `kg_related` helps resolve ambiguous outcomes

One of the strongest design choices in this project is that its resolution logic uses explicit outcomes. That may sound like a small implementation detail, but it changes how you build reliable pipelines. A system that always returns something invites silent errors. A system that can say `HOLD`, `AMBIGUOUS`, or `NO_CANDIDATE` is asking for discipline.

`kg_related` is often the best tool to interrogate those uncertain states.

If a record lands in `AMBIGUOUS`, it usually means the top candidates are too close on the currently available evidence. Looking at each candidate’s neighboring entities can expose a clean split. One candidate may be tied to the right university or coauthor. Another may belong to a different country or field entirely. Even when you do not automate the final decision, the human reviewer gets a far clearer picture.

If a record lands in `HOLD`, the issue is often incomplete or conflicting evidence. Related entities can show whether there is at least a coherent local neighborhood worth preserving for later review. Sometimes that keeps you from discarding a candidate too early. Sometimes it reveals that the apparent match is only a label collision.

If a record lands in NO_CANDIDATE, kg_related is usually not the first next step, because you need something to relate from. But after refining search terms or discovering a near miss, related context may explain why no safe candidate exists yet. The local record may refer to a niche entity that is absent, underdeveloped, or named differently in Wikidata.

A practical workflow that keeps errors down

When I have to link a messy local dataset, I avoid trying to be clever too early. The most robust flow is compact:

- use kg_search to get a bounded set of candidate entities
- inspect the strongest candidates with kg_entity, focusing on a few high-signal facts
- use kg_related only when candidate context is the missing piece
- if uncertainty remains, keep the record in HOLD or AMBIGUOUS rather than forcing a match
- use optional Google cross-checks as concordance support, never as sole proof

That sequence respects the way the server is designed. It also scales. In batch work, the expensive part is not usually one extra API call. It is the cleanup cost from bad identity decisions. kg_related saves time when it prevents false confidence.

Cases where kg_related is the wrong tool

Not every record needs graph context. Sometimes people overuse relatedness because it feels richer. Richness is not always efficiency.

If you already have a distinctive identifier in your local data that maps cleanly to a single Wikidata entity, kg_related may be unnecessary. If you need selected facts with ranks, qualifiers, or references for a known entity, go straight to kg_entity. If the problem is that search is too broad or your label variant is poor, fix the candidate generation step first.

There are also records where “relatedness” creates more noise than signal. Broad concepts, mass-market works, or entities with huge public footprints can have many plausible neighbors. In those cases, a very specific property from kg_entity may carry more weight than a general sense of connectedness.

I would skip kg_related in these situations:

- the entity is already confidently identified and you only need facts from its own record
- the local record lacks any relational clues, so there is nothing meaningful to compare against
- your main issue is poor candidate search, not candidate validation
- the entity is so broad that neighboring nodes add volume without discrimination
- your review process cannot inspect contextual evidence and needs a simpler rule

That last point matters more than many teams admit. If your downstream process cannot consume contextual evidence, kg_related can become decorative. It is best used when either a person or a well-designed rule can act on what the related entities imply.

Reading relatedness without overreading it

This is where judgment matters. A relationship in the graph can be highly informative, mildly helpful, or nearly useless depending on the domain. In biographical matching, affiliation, office, employer, collaborator, and place links can be decisive. In cultural works, a cast or contributor connection might help, but only if your local record already hints at that context. In geography, neighboring administrative relationships may settle the matter quickly.

The trap is treating any plausible connection as confirmation. The project's stance on evidence should keep you honest here. It is designed for inspectable evidence and explicit uncertainty. Relatedness should reduce uncertainty by aligning with known local clues. If it merely looks interesting, it has not done enough work.

A simple mental test helps. Ask whether the relationship would persuade a cautious reviewer who had not seen the match before. "This candidate is connected to the same university named in the local record" is persuasive. "This candidate is connected to other academics" usually is not.

kg_related and human review queues

One of the best uses for kg_related is not fully automated resolution but triage. When you have a queue of records that need review, context can dramatically cut handling time. A reviewer who sees a likely candidate plus a small amount of well-chosen neighborhood evidence can make a decision much faster than a reviewer staring at names alone.

That fits the project's overall model. The server is read-only. It does not edit Wikidata, Google, or user data. It is also not official Wikimedia or Google software, and not an export of the Google Knowledge Graph. Its job is to help an MCP client explore and compare evidence, not to declare truth by fiat.

In that environment, kg_related is especially valuable because it turns a thin candidate into a more intelligible one. You are not replacing review. You are making review more grounded.

Where this sits in the broader MCP landscape

Wikidata's own MCP documentation frames the broader purpose clearly: standardized tools for LLMs to explore and query Wikidata programmatically through the Wikidata API and Query Service. The Wikidata + Google Knowledge Graph MCP server takes a narrower and more operational approach. It focuses on search, selected facts, bounded candidate sets, and entity resolution with explicit uncertainty. That is why the tool set feels deliberate.

Within that smaller [Wikidata MCP](#) tool set, kg_related fills a distinct role. It is the contextual bridge between retrieval and resolution. It helps an agent move from "I found something" to "I understand why this is, or is not, the right thing."

For MCP for google knowledge graph and wikidata work, that role is easy to underestimate. Many pipelines obsess over matching algorithms and ignore evidence presentation. But if you have ever cleaned up a false-merge problem in a catalog, a CRM, or a research corpus, you learn quickly that the cost of weak evidence compounds. A bounded, inspectable relatedness step is often cheaper than retroactive correction.

The simplest rule of thumb

Use kg_related when the identity question depends on the company an entity keeps.

That sounds almost too simple, but it captures the practical boundary. If your local record can be validated by the entity's immediate facts alone, use kg_entity. If you do not yet have a plausible entity, use kg_search

or `kg_resolve`. If the deciding evidence lives in affiliations, collaborators, places, organizations, or other neighboring entities, use `kg_related`.

Used that way, it complements the rest of the server instead of duplicating it. It gives you just enough graph context to make better decisions, while staying aligned with the project's core principles: bounded search, inspectable evidence, and explicit uncertainty when the data does not support a clean match.