University of Otago
ŌTĀKOU WHAKAIHU WAKA

Te Kura Rorohiko
School of Computing

Are LLMs and the Model Context Protocol sufficient for automating software web-based information processing?

Stephen Cranefield

School of Computing,
University of Otago,
Dunedin, New Zealand

A motivating scenario for the Semantic Web

(Berners-Lee et al. Scientific American, 2001)

- Lucy and Peter need to arrange specialist medical treatment for their mother.
- Tasks:
  - Find a provider of the treatment that …
  - is well rated,
  - close to her home,
  - approved by her health insurance company,
  - and has availability at times when either Lucy or Peter are available.
- Vision: This would be enabled by *agents* and *ontologies*

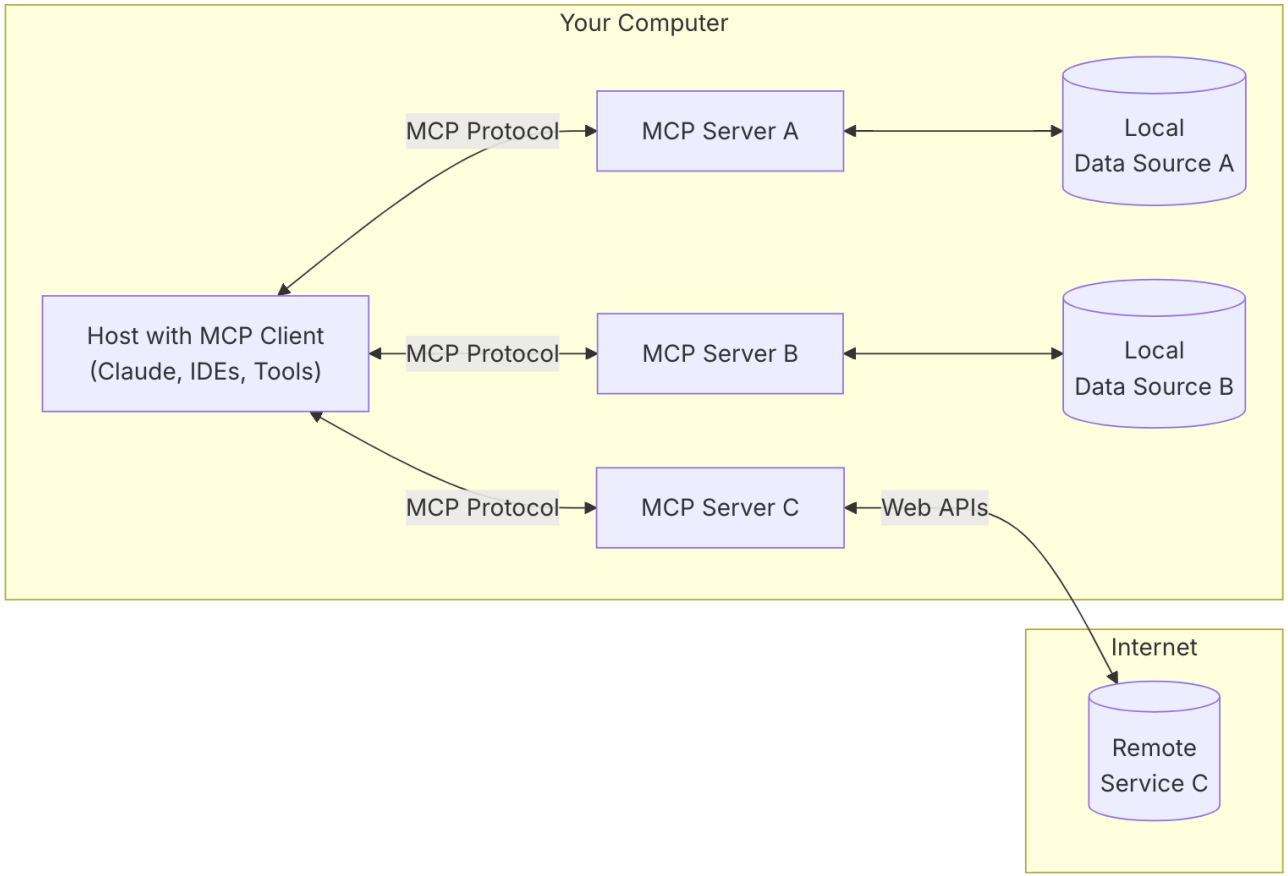
See illustration of scenario at https://static.scientificamerican.com/sciam/cache/file/394EDA92-D03F-4110-B5AA4465CE486800.pdf#page=31

Can we automate this at last using LLM-powered agents and MCP?

- The model context protocol (MCP)

  "… an open-source standard for connecting AI applications to external systems"

  https://modelcontextprotocol.io/docs/getting-started/



MCP diagram from https://github.com/NirDiamant/agents-towards-production/blob/main/tutorials/agent-with-mcp/mcp-tutorial.ipynb. Author: Nir Diamant.
Licenced for non-commercial use

My attempt

- Preliminary study
- Simplified scenario
  - Find physiotherapists in Dunedin, NZ
  - that are “FirstChoice” providers for the nib insurance company.
  - Choose the one that is closest to the user’s address
  - Ask user to contact this provider and enter the first three available appointment times in the chat
    (a proxy for the LLM doing this)
  - Check the user’s calendar to see at which appointment time they are free
    (simulated via chat)

- **MCP Host:** LM Studio on MacOS
- **LLM:** qwen2.5-7b-instruct running locally
- **MCP servers:** Fetch, MCP Http, Mapbox

Eventual success

Iterative prompt development

- I had to provide the required web sites
- … and guidance on how to use them
- … as well as some tool-choice tips

- https://nzdirectory.co.nz/search?query=<BUSINESS_SEARCH_TERM>&region=<REGION>&city=<CITY>. This is the search URL pattern for a directory of businesses in New Zealand. For <BUSINESS_SEARCH_TERM> you can use a sufficiently discriminating prefix of the business or service type. Given a <CITY> you MUST choose <REGION> to be whichever of the following New Zealand regions contains the city: Northland, Auckland, Waikato, Bay of Plenty, Gisborne, Hawke's Bay, Taranaki, Manawatu-Whanganui, Wellington, Tasman, Nelson, Marlborough, West Coast, Canterbury, Otago and Southland.
  Use your background knowledge to choose the correct region for the city. You must convert the region and city to lowercase when including them in the search URL. Performing a GET on the search URL returns an HTML page for human viewing, so use a tool that converts the result to a more concise format like Markdown.
- https://www.nib.co.nz/find-a-provider/api/recommendations?searchTerm=<SEARCH_TERM> is the search URL for the insurance company nib's FirstChoice providers. The search term can be a sufficiently distinguishing prefix of the medical speciality of interest (in lowercase). The results are in JSON, so obtain the results using a tool that does not modify the format. Examine the JSON to find the details for the returned providers.

Conclusion: Solutions are needed for …

- Tool documentation, selection and calling
  - I chose the MCP servers to provide to the LLM use
  - I had to tell the LLM *when* to use *which* HTTP tool
    - E.g. because one converts HTML to Markdown and one doesn't.
  - MCP provides rather limited tool information to LLMs
    - A "human-readable description of functionality"
    - Procedure names
    - Input schemas
    - No output schema!
  - Example issue:
    The Mapbox tool that makes distance calculations gives the LLM no information about measurement units

- https://nzdirectory.co.nz/search?query=<BUSINESS_SEARCH_TERM>&region=<REGION>&city=<CITY>. This is the search URL pattern for a directory of businesses in New Zealand. For <BUSINESS_SEARCH_TERM> you can use a sufficiently discriminating prefix of the business or service type. Given a <CITY> you MUST choose <REGION> to be whichever of the following New Zealand regions contains the city: Northland, Auckland, Waikato, Bay of Plenty, Gisborne, Hawke's Bay, Taranaki, Manawatu-Whanganui, Wellington, Tasman, Nelson, Marlborough, West Coast, Canterbury, Otago and Southland.
  Use your background knowledge to choose the correct region for the city. You must convert the region and city to lowercase when including them in the search URL. Performing a GET on the search URL returns an HTML page for human viewing, so use a tool that converts the result to a more concise format like Markdown.
- https://www.nib.co.nz/find-a-provider/api/recommendations?searchTerm=<SEARCH_TERM> is the search URL for the insurance company nib's FirstChoice providers. The search term can be a sufficiently distinguishing prefix of the medical speciality of interest (in lowercase). The results are in JSON, so obtain the results using a tool that does not modify the format. Examine the JSON to find the details for the returned providers.

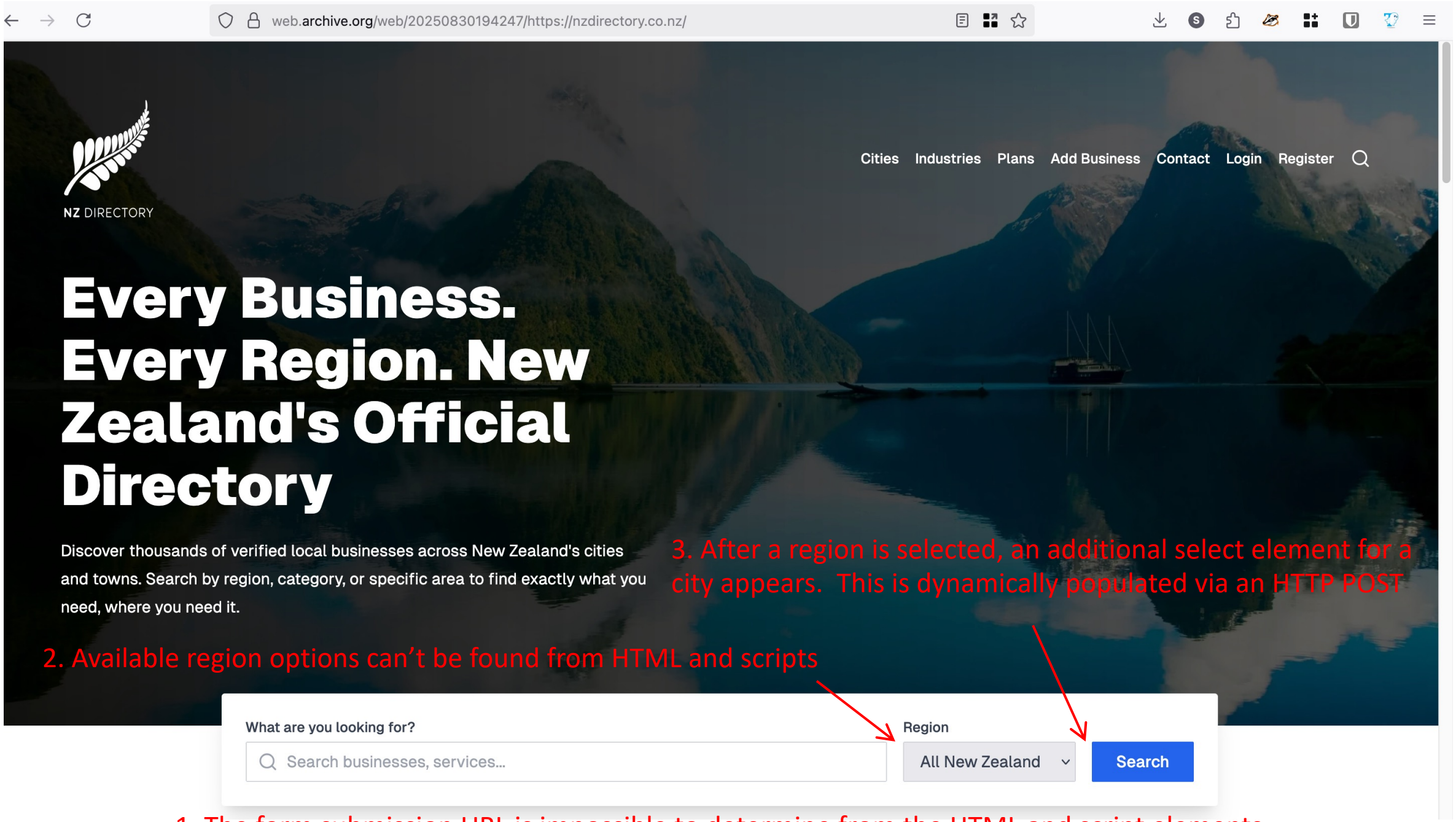
Conclusion: Solutions are needed for …

- Service discovery and selection
  - How will LLMs discover web sites that can answer a user's query? Some options:
    - A standard for including LLM-oriented service descriptions in web indexes (in natural language for LLM consumption or maybe WoTs "thing descriptions"*)
    - A parallel ecosystem of web site registries and catalogues specifically for the use of LLMs.
    - Community-provided comments and ratings and web sites accessibility to LLMs, plus usage tips.

* Are these any more likely than ontologies to gain widespread acceptance?

Conclusion: Solutions are needed for …

- Navigating web sites from a single entry point
  - i.e. given a top-level URL can an agent navigate to the web pages relevant to its goals.
  - LLM agents will need this more than people.
  - Modern web development frameworks do not seem to encourage good practice
  - Example issue: next slide


web.archive.org/web/20250830194247/https://nzdirectory.co.nz/
NZ DIRECTORY
Cities
Industries
Plans
Add Business
Contact
Login
Register
Every Business. Every Region. New Zealand's Official Directory
Discover thousands of verified local businesses across New Zealand's cities and towns. Search by region, category, or specific area to find exactly what you need, where you need it.
3. After a region is selected, an additional select element for a city appears. This is dynamically populated via an HTTP POST
2. Available region options can't be found from HTML and scripts
What are you looking for?
Search businesses, services...
Region
All New Zealand
Search
1. The form submission URL is impossible to determine from the HTML and script elements

Conclusion: Solutions are needed for …

- Reasoning at a goal or task level