

Please send some version of this demo to [akash@euprime.org](mailto:akash@euprime.org). There are more than 1000 applications. As a result, sending a working demo of the following would help us in adjudicating your candidature. Even if you do part of the objective, that is fine too. Just have something along the lines described below. Thanks a ton. You can follow us on our youtube channel <https://www.youtube.com/watch?v=aXNxOlab7Yw>

I want to create a web agent/crawler that would help me with the following. Think and act like a business developer that crawls the web for relevant information to output very well qualified leads. The tool should then be able to, based on the criteria we provided, prioritize or rank the probability of these people wanting to work with us.

3D in-vitro models helping researchers to design new therapies.

Stage 1: Identification (The Input)

It scans target profiles based on your criteria (e.g., Director of Toxicology, Head of Preclinical Safety) from LinkedIn Search or Sales Navigator.

Data Source: LinkedIn, PubMed (for authors of recent relevant papers),  
and Conference attendee lists (e.g., SOT - Society of Toxicology).

Stage 2: Enrichment (The Data Gathering)

o For each identified person, the tool queries external databases to find:

Contact Info: Business email (e.g., @pfizer.com) and phone numbers.

Location Data: It distinguishes between the Person's Location (e.g., remote in Colorado) and the Company HQ (e.g., Cambridge, MA).

Stage 3: Ranking (The Probability Engine)

o It applies "Propensity to Buy" score to each person based on weighted criteria (detailed below).

## 2. The Probability Ranking Logic

To prioritize who wants to work with 3D-invitro models for therapies, the tool would assign a score (0-100) based on these weighted signals.

Signal

Category Criteria Example Scoring

Weight

Title contains

Role Fit

Toxicology, Safety, Hepatic, 3D - High(+30)

Company Intent

Company recently raised Series A/B funding (cash to spend), High (+20)

Technographic

Company already uses similar tech (e.g., in vitro models), Medium(+15)

open to New Approach Methodologies (or NAMs), Medium(+10)

Medium

(+15)

Location Located in a Hub (Boston/Cambridge, Bay Area, Basel,  
UK Golden Triangle)

Medium

(+10)

Scientific

Intent

Published a paper on Title contains  
(Drug-Induced Liver Injury) in  
last 2 years

Very High

(+40)

Result: A list of 500 people is instantly re-ordered.

Junior Scientist at a non-funded

startup might score 15/100, while a Director of Safety Assessment at a Series B  
biotech in Cambridge, MA who just published on liver toxicity would score 95/100.

## The Output: The Lead Generation Dashboard

The final output would be a dynamic, searchable table (exportable to Excel/CSV) with the following columns:

Rank Probability Name Title Company Location HQ Email LinkedIn Action

Search Functionality: You can type Boston or Oncology to filter this list

instantly.

Location Split: The table explicitly shows if the person is remote (e.g., Living in

Texas, HQ in Boston), which helps BD decide when to call or if a physical meeting is possible.

## 4. How to Build/Deploy This

Since you want this eventually to be a there are two paths:

Use a tool like Clay or Apollo.io.

o Clay is specifically designed to do exactly this: You feed it a LinkedIn search URL, it finds the emails, maps the locations, uses AI to read their company description to see if they drug discovery, assigns a score, and puts it all in a Google Sheet/Dashboard.

### Path B: Custom Dashboard

o A simple web app (Python/Streamlit) that connects to the LinkedIn API (or a provider like Proxycurl) to fetch data and renders the table exactly as

described above.

Patch C: You can use your own new methodology

what are some of the main types of websites that the web agent needs to search through to procure the information to be efficient?

To identify, enrich, and rank the best leads, the web agent needs to systematically harvest data from five specific categories of websites. This approach is what differentiates a generic lead list from the high-probability tool you are building.

#### 1. Professional Social Networks

This is your baseline for identification and employment data.

Primary Source: LinkedIn (and Sales Navigator).

Secondary Source: Xing (relevant since this technology has a strong European footprint/HQ).

What the Agent looks for:

- o Job Titles: Director of Toxicology, Head of Safety Assessment, VP

Preclinical

- o Location: Distinguishing between Home (Remote) vs. Office

- o Tenure: How long they have been in the role (new hires are often looking to buy new tech).

#### 2. Scientific Publication Databases

Primary Sources: PubMed, Google Scholar, bioRxiv.

What the Agent looks for:

- o Keywords: (Drug-Induced Liver Injury), 3D cell culture, Organ-on-chip, Hepatic spheroids, Investigative Toxicology
- o Recency: Papers published in the last 12-24 months (indicates active research).
- o Authorship: Focus on the Corresponding Author (usually the PI or budget holder) or the First Author (often the user/influencer).

### 3. Biomedical Conference Event Sites (The Active Market)

People attending these events are actively looking for solutions and networking.

Primary Sources: SOT (Society of Toxicology), AACR (Cancer Research), ISSX (Xenobiotics), ACT (College of Toxicology).

- o Exhibitor Lists: Which competitive or partner companies are attending?
- o Poster/Abstract Presenters: Scientists presenting data on liver models or safety assessment.
- o Agenda Speakers: Key Opinion Leaders (KOLs) in the space.

### 4. Business & Funding Intelligence (The Budget)

This determines if the target actually has the cash to buy premium solutions.

Primary Sources: Crunchbase, PitchBook, FierceBiotech, Biopharm Guy.

What the Agent looks for:

- o Funding News: Series A, Series B, IPO, New Grant. (e.g., A biotech raising \$50M is a prime target).
- o Partnerships: News about a company entering a & Phase 1 Safety&quot; trial (indicates they are past discovery, but might need help with the next pipeline asset).

## 5. Public Grant Databases (The Academic Budget)

For targeting academic or government researchers who aren't on Crunchbase.

Primary Sources: NIH RePORTER (USA), CORDIS (EU).

What the Agent looks for:

- o Grant Awards: grants containing keywords like  
& Liver Toxicity or & 3D Models
- o Grant Expiry: Finding grants that just started vs.  
those ending.