

Data Is King

Why data ownership and quality set the ceiling on everything AI can responsibly do for you.

Every AI conversation eventually arrives at the same place. Not the model, not the prompt. The data. What you own, how good it is, and what you are allowed to do with it decide the ceiling on everything else.

Ownership comes first. A capability built on data you rent can be priced out from under you, and a capability built on data you hold compounds. When I needed a nationwide picture of which districts existed and how to reach them, the value was never the clever query. It was having a clean, current, verified registry that was mine.

Quality is the next constraint, and it is unglamorous work. Duplicates, stale records, fields that mean different things in different rows, a state that imports under one identifier while every other state uses another. AI will happily reason on top of all of it and produce an answer with no hesitation at all. The output looks the same whether the input was right or not.

So instrument the inputs. Know where each field came from, when it was last checked, and who verified it. A record that carries its provenance is worth several that do not.

Permission is the third piece. Consent, retention, minimization, and access control are not legal decoration. In my world that means a parent knows exactly what is collected about their child, what is shown publicly, and what is not. The rules shape the product, and they should, because a platform that mishandles that data does not get a second chance.

Do this well and something useful happens. Your data becomes the thing no competitor can copy by buying the same tools you bought. They can run the same model. They cannot run your record of what is actually true.

The rule: the model is rented, the data is owned. Invest accordingly.