

Where We Are Now

A plain-English summary of what Cleopas has become, why it is better, and what remains before MVP.

THE SIMPLEST VERSION

Cleopas does not tell people which expert to believe. It breaks a difficult question into smaller claims, shows the full evidence landscape, tests how much each fact should matter, and then explains the result in ordinary language. The engine became more disciplined so the conversation can become simpler.

What changed? We moved from claim + citation + tier to an evidence landscape: many smaller claims, different evidence categories, source limits, alternative explanations and human motive.	Is it better? Yes. It is harder for one citation, one expert or one persuasive argument to control the conclusion.
---	--

What did the research teach us? T1-T4 remains useful, but the tier is the result. The real product is showing why the evidence earns that result.	What must MVP prove? That Cleopas can hold a natural conversation while every historical sentence remains inside a traceable evidence boundary.
---	---

Two common arguments, handled differently

"Roman law says crucified bodies were left exposed, so Jesus was not buried."	Cleopas: Roman practice matters, but first identify the exact document, its date, its exceptions, whether it records law or practice, whether officials followed it, and how directly it applies to Pilate. Then place it beside the other evidence categories.
"The tomb was empty, so Jesus rose from the dead."	Cleopas: An empty tomb, if historical, establishes a missing body from a place. It does not identify the cause. Theft, relocation, error, later narrative development and resurrection remain separate explanations that require separate evidence.

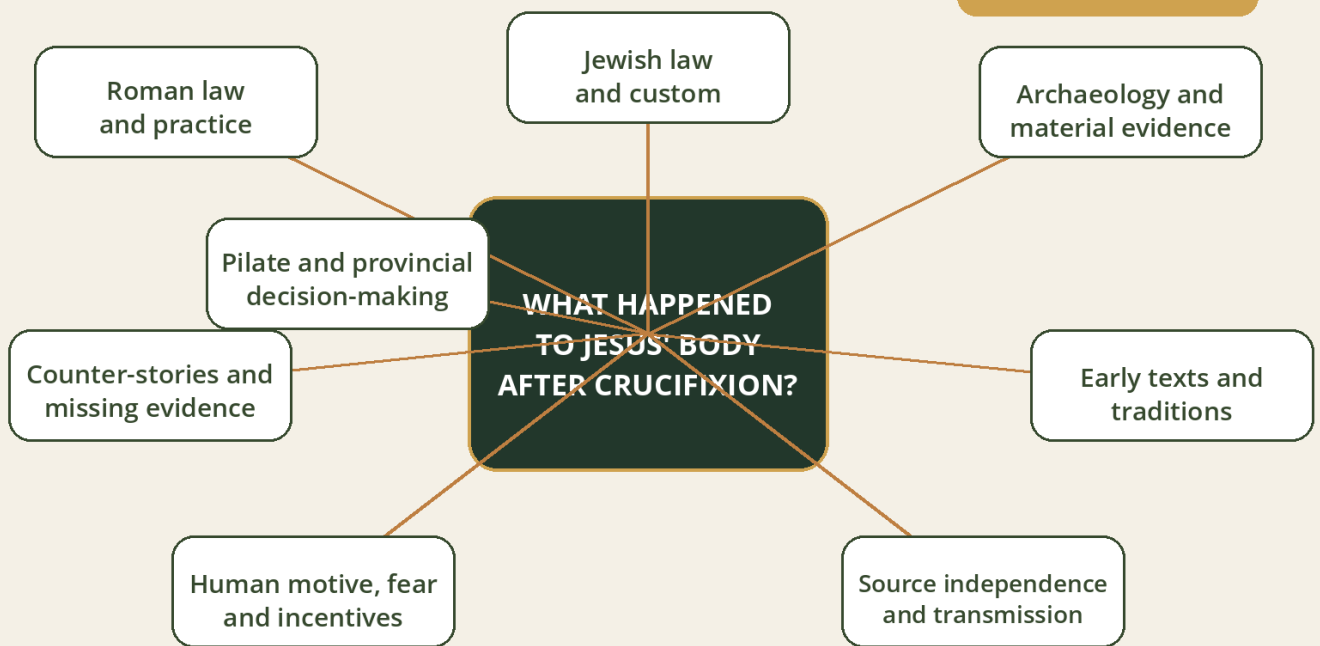
The Evidence Landscape

The depth is not there to overwhelm the user. It is there to stop one persuasive point from pretending to be the whole case.

A person may say, "Roman law says X, therefore I believe X." Cleopas does not begin by arguing. It shows that the conclusion sits inside a wider landscape. A single event may involve 30 or more considerations across eight major categories, and those considerations do not all perform the same job.

ONE QUESTION. EIGHT EVIDENCE CATEGORIES.

Cleopas does not let one strongly argued point become the whole case.



The user can still give Roman practice more weight. But they now know what else they are choosing to discount.

The core product insight

CLEOPAS DOES NOT TELL PEOPLE WHAT TO THINK.

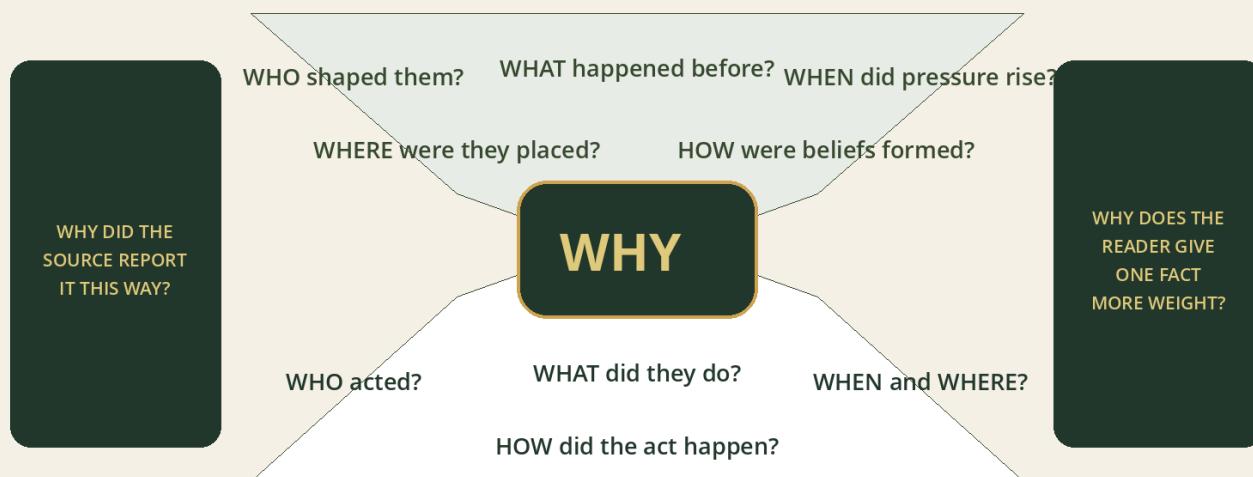
It shows them how much there is to think about. It can surface agreement, disagreement, missing evidence and the assumptions that cause a grade to move without declaring that one expert, one tradition or one worldview owns the answer.

Why Matters

Who, what, when, where and how describe an event. Why connects the event to the people, pressures and beliefs that caused it.

THE WHY HOURGLASS

The visible event sits below a deeper set of causes, pressures and beliefs.



Motive can change plausibility. It cannot become proof unless the evidence earns it.

Three different “why” questions

Why did the actor act?	Why did the source report it?	Why does the reader weight it?
Pilate may have been harsh, but harshness alone does not decide whether he would grant burial. Cleopas tests incentives, costs, public order, local pressure and discretion.	A source can preserve memory, inherit tradition, simplify, shape, persuade or invent. Those are separate possibilities, not a binary choice between truth and lie.	A user may trust legal evidence, distrust religious sources or carry prior experiences. Cleopas can expose that weighting without attacking the person.

The discipline

- A stated motive can be quoted. An inferred motive must be graded. A plausible incentive can move probability. A speculative story must remain speculation.
- Sequential claims do not inherit the meaning of what follows. Burial does not prove an empty tomb. An empty tomb does not prove resurrection.

“Roman Law Says X”

Even evidence that looks neutral, official or non-religious must pass the same audit.

THE LOW-CONTESTATION BIAS

Some claims look unusually secure because they do not threaten anyone’s identity or worldview. That may mean they have received less adversarial scrutiny, not that their evidence is automatically better. Internally, we call this the Dinosaur Test: would we question this source just as deeply if it affected a belief people cared about?

SOURCE AUDIT • Which document exactly? • When was it written or compiled? • What earlier material does it preserve? • How was it transmitted and preserved? • Is it legislation, juristic opinion, custom or a report of an event?	APPLICATION AUDIT • What exceptions and permissions appear in the text? • Who held discretion? • Do recorded officials follow or breach the rule? • Does it apply to Judea under Pilate? • How much should it matter beside the other evidence?
---	--

What the Roman example already shows

- The famous Digest passage is not a simple first-century Judean statute. The Digest was compiled under Justinian in the sixth century from extracts of earlier jurists, many from the late second and early third centuries.
- Digest 48.24 itself preserves a general preference for returning bodies, a request-and-permission process, official discretion and a stated treason exception. The legal material is not a single universal rule of non-burial.
- Roman writers also preserve cases where provincial officials ignored, stretched or manipulated legal protections. Law, enforcement and individual behaviour must be tested separately.
- Therefore: “Roman law says X” may be a strong input. It is not automatically a decisive conclusion about what Pilate did with one body in one circumstance.

Source note: Justinian, Digest 48.24 (Ulpian and Paulus); J. F. Gardner, “Slavery and Roman Law,” Cambridge World History of Slavery; “Extreme Justice,” Britannia 52 (2021); Cicero, Against Verres 2.5. These sources illustrate the method and do not themselves settle Jesus’ burial.

A Contested Event Contains Mixed Grades

The parent question should not be forced into one oversized claim. Break it apart until the evidence can be graded honestly.

For every atomic claim, Cleopas separates two questions:

1. How sure are we this is true? This is the T1-T4 confidence grade.	2. How much should it matter here? This is relevance and impact on the larger question. A very certain contextual fact may still have only moderate effect on one event.		
Atomic claim	Confidence	Relevance	Role in the larger question
Jesus was executed under Pontius Pilate	T1	High	Event foundation
Roman crucifixion could include prolonged exposure	T1/T2	Moderate	Raises a non-burial prior
Roman authorities could permit release for burial	T2	Moderate	Keeps burial scenarios open
Jewish tradition strongly valued prompt burial	T1/T2	Moderate	Context, not Roman compliance
A crucified person in first-century Jerusalem received burial	T1	Moderate	Establishes occurrence, not frequency
An early Christian burial tradition existed	T1	High	Early attestation, not every detail
Jesus received some form of burial	T2?	High	Current cautious working view
Joseph of Arimathea arranged the burial	T3	High detail	Serious but contested
An identifiable tomb was later found empty	T3 / research	High	Does not identify the cause
A missing body was caused by resurrection	Outside history	Ultimate	Requires more than ordinary historical grading

Illustrative only: final grades remain provisional until the current three-reviewer cycle and open research branches are complete.

NO REVIEWER IS THE ANSWER.
Experts identify facts, challenge source relationships and expose hidden assumptions. Their confidence and passion do not transfer directly into Cleopas. The same method must decide how every dimension affects the final grade.

What MVP Must Prove Next

The next milestone is not solving every branch. It is proving that the governed method works on one difficult event.

REQUIRED FOR MVP

- A small reviewed set of atomic claims and source passages.
- AI first-pass extraction of claims, sources, dates, evidence type and candidate limits.
- Human approval for source independence, motive, weighting, confidence and maximum wording.
- Runtime retrieval that surfaces support and counterweight without fake 50/50 balance.
- Answers that remain inside the same evidence boundary in CRUISE, OPEN HOOD and SPAR modes.
- A measurable tagging benchmark rather than an assumed 70/30 automation split.

OPEN RESEARCH MAP

- Pilate: why might he grant or refuse burial?
- The missing body and earliest counter-story.
- Roman legal sources: date, transmission, enforcement and Judean application.
- Causal and motive model: actor, source and user.
- Arguments from silence and counterfactual disconfirmation strength.
- Claim-specific Gospel independence, especially John.
- Competing burial scenarios and which facts survive each one.
- Website demos that expose depth without sounding technical.

What the tagging hypothesis now means

The stronger rule set should make AI-assisted structuring easier because the model receives smaller, clearer questions. It can likely identify claims, quotations, dates, named sources and basic evidence roles at high rates. The difficult work is not extraction. It is deciding source independence, motive, relevance, decisiveness, arguments from silence and final confidence. Those consequential judgments remain human-approved until the MVP measures where automation is genuinely reliable.

THE PRODUCT PROMISE

Cleopas does not win by sounding certain. It earns attention by showing the full landscape, questioning every source by the same standard, and explaining why a conclusion deserves exactly the confidence it receives - no more and no less.