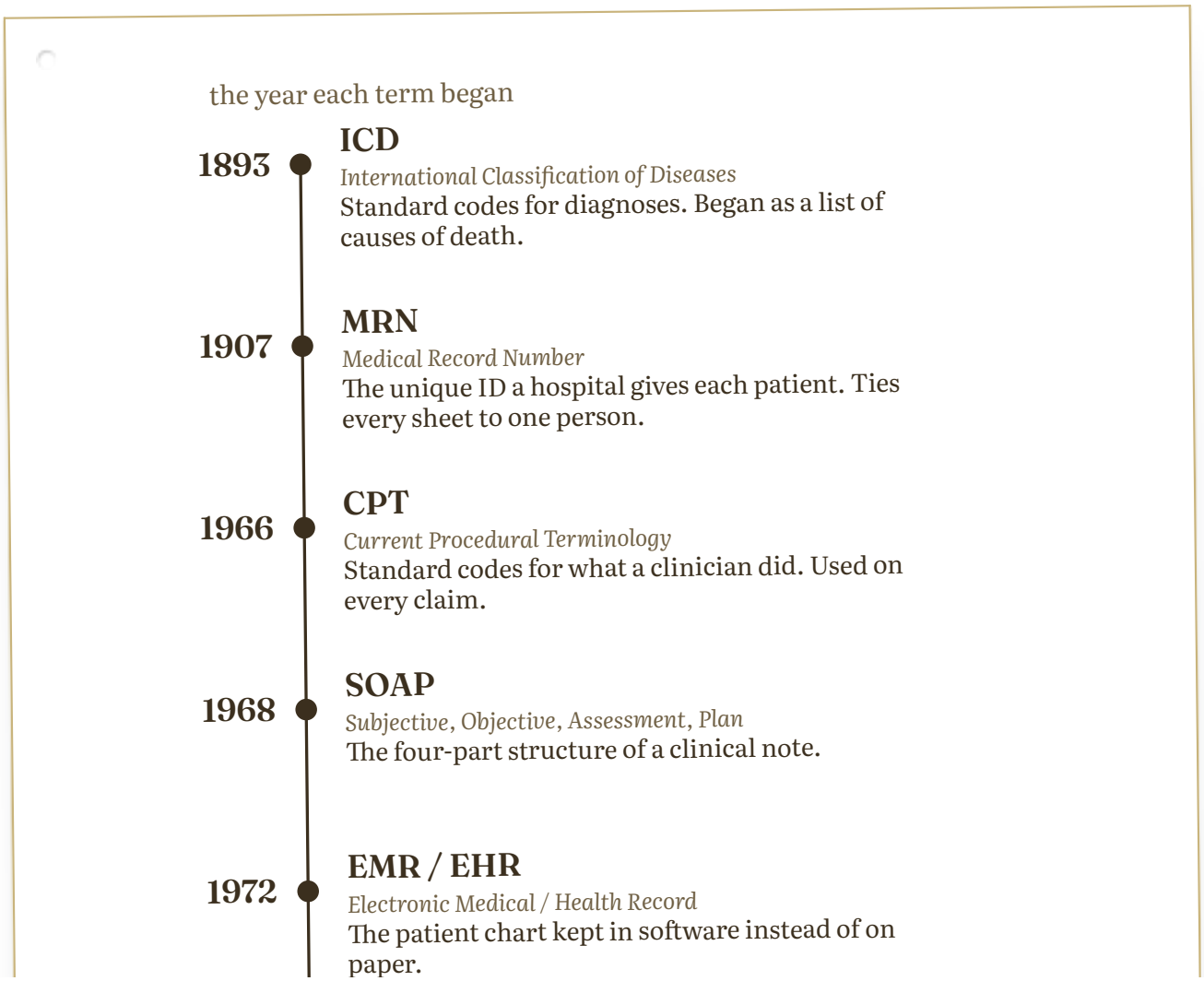
Medical record acronyms: what each one means and why it exists

Fourteen terms you will meet in any healthcare product. For each: what it is, where it came from, and why it matters to the software.

September 2026, about a ten minute read



1987 — HL7
Health Level Seven
The messaging standard hospital systems use to talk to each other.

1996 — HIPAA
Health Insurance Portability and Accountability Act
The US law whose Privacy and Security Rules govern health data.

2000 — PHI
Protected Health Information
Any health data that can identify a patient. What HIPAA protects.

2000 — Safe Harbor
HIPAA de-identification method
Remove 18 listed identifiers and the data is no longer PHI.

2003 — BAA
Business Associate Agreement
The contract required with any vendor that handles PHI.

2007 — NPI
National Provider Identifier
A single public ten-digit number for every clinician and organization.

2009 — HITECH
Health Information Technology for Economic and Clinical Health Act
Paid hospitals to adopt EHRs and made vendors directly liable for breaches.

2011 — FHIR
Fast Healthcare Interoperability Resources
A web API standard for reading and writing health records.

2015 — ICD-10-CM
ICD, 10th revision, Clinical Modification
The diagnosis code set the US switched to on October 1, 2015.

Each term with the year it started, what it stands for, and what it does.

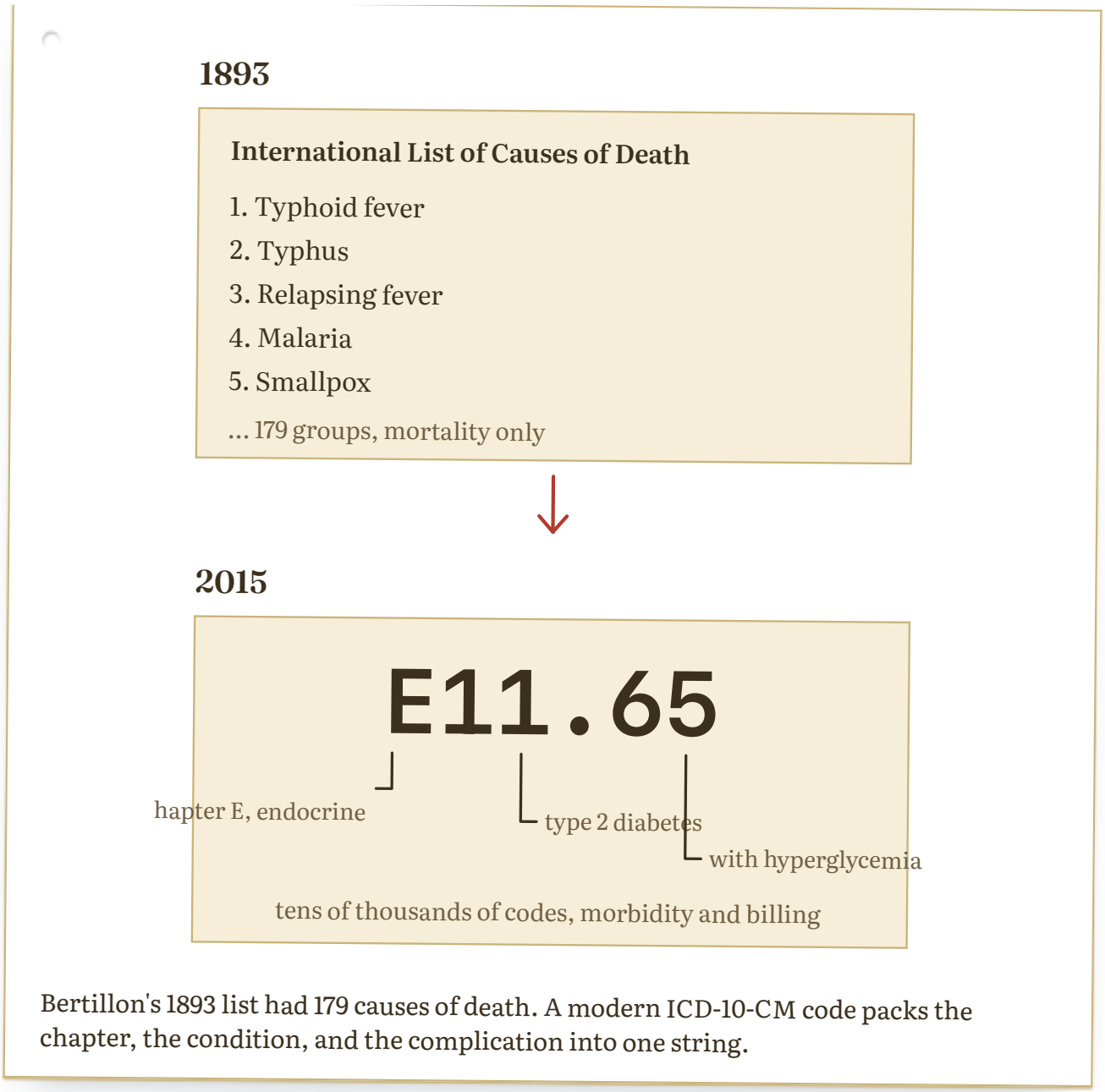
These are listed in the order they were created. The technology terms (scanning, text extraction, AI models) are left out on purpose. Everything here would still exist if computers didn't.

ICD

International Classification of Diseases, 1893

What it is: the standard code set for diagnoses. "E11.65" means type 2 diabetes with high blood sugar. Every diagnosis in a chart or on a claim is an ICD code.

Where it came from: a statistician in Paris, Jacques Bertillon, was asked to make a list of causes of death that countries could compare. It was adopted in 1893 and the US started using it in 1898, with an update every ten years. In 1948 the World Health Organization took it over and expanded it from causes of death to all diseases. The US used its own version, ICD-9-CM, for far longer than most countries and only switched to ICD-10-CM on October 1, 2015.



Bertillon's 1893 list had 179 causes of death. A modern ICD-10-CM code packs the chapter, the condition, and the complication into one string.

> **Why it matters:** a diagnosis code is structured data you can rely on, unlike free text. It is also a clue to what is in the note.

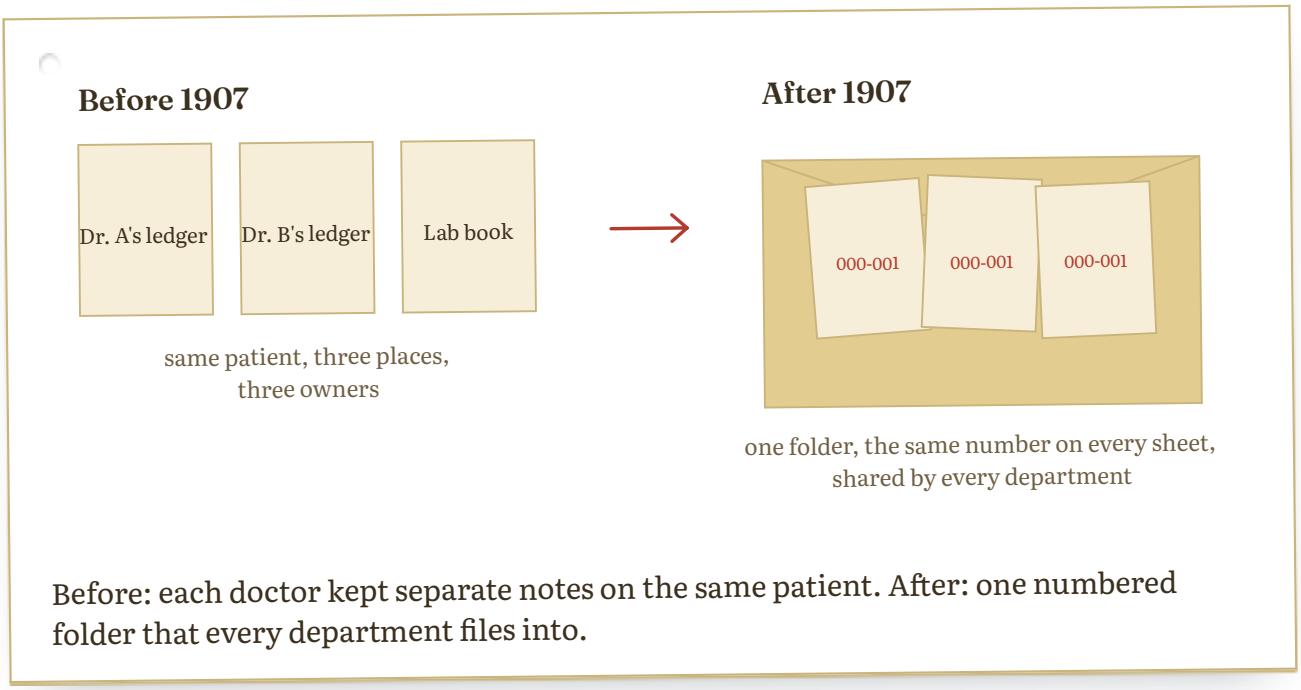
MRN

Medical Record Number, 1907

What it is: the unique ID a hospital assigns to each patient. Every document about that patient carries it.

Where it came from: until 1907, doctors at the Mayo Clinic each kept

their own notebook. A patient who saw two doctors had two unrelated records. Henry Plummer, a physician who had wanted to be an engineer, replaced this with one folder per patient, kept in a central file, with a number stamped on every sheet. Patient number 000-001 was entered on July 19, 1907. Every hospital record system since has copied the idea.



Before: each doctor kept separate notes on the same patient. After: one numbered folder that every department files into.

> **Why it matters:** an MRN identifies a patient directly, so it counts as protected information and has to be removed when data is de-identified (see Safe Harbor below).

CPT

Current Procedural Terminology, 1966

What it is: the standard code set for procedures and services. ICD says what was wrong; CPT says what the clinician did about it.

Where it came from: the American Medical Association published the first edition in 1966 so that a service meant the same thing to every insurer. The AMA still owns and licenses it.

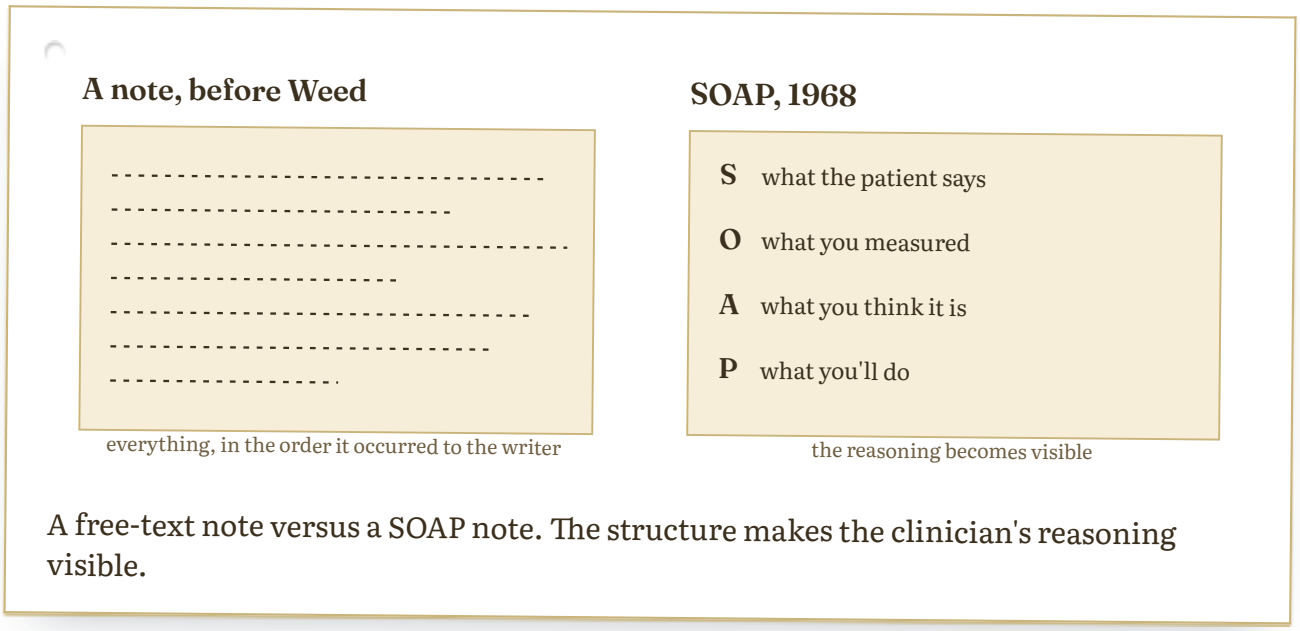
> **Why it matters:** ICD plus CPT is the language of billing. In a hospital, "coding" means assigning these codes, not writing software.

SOAP

Subjective, Objective, Assessment, Plan, 1968

What it is: the standard structure of a clinical note. What the patient reports, what the clinician measured, what the clinician thinks is going on, and what happens next.

Where it came from: Lawrence Weed, a physician in Vermont, published a paper in 1968 arguing that medical notes were so disorganized you could not tell whether the doctor had reasoned clearly. He proposed organizing the chart around a numbered list of problems, with a four-part note for each one. The full system was too much work and mostly did not catch on. The four-part note did.



A free-text note versus a SOAP note. The structure makes the clinician's reasoning visible.

> **Why it matters:** most notes you will parse follow this shape, even when the headings are missing. Knowing it helps you find the assessment and the plan.

EMR and EHR

Electronic Medical Record, Electronic Health Record, 1972

What it is: the patient chart kept in software. The two terms are used interchangeably. Strictly, a medical record is one organization's chart and a health record is meant to follow the patient across organizations.

Where it came from: several groups built early versions in the late 1960s and early 1970s. The one most often credited is the Regenstrief Institute in Indianapolis, which in 1972 built a system that pulled lab and pharmacy data into one record automatically. The Veterans Administration started its own system around the same time. Adoption stayed low for decades because computing was expensive; in the late 1960s a megabyte of memory cost about a thousand dollars.

> **Why it matters:** the EHR is where the data lives. Reading from it and writing back to it is most of the integration work.

HL7

Health Level Seven, 1987

What it is: the messaging standard hospital systems use to talk to each other. An HL7 message says things like "this patient was admitted" or "this lab result is ready."

Where it came from: once a hospital had a lab system, a pharmacy system, and an admissions system from different vendors, they needed a shared message format. A group of volunteers founded HL7 in 1987 to define one. The "seven" refers to the application layer of the network stack: the standard is about what messages mean, not how they travel. HL7 version 2, from that era, is still the most common interface in healthcare. Version 3 was more rigorous and mostly

went unused.

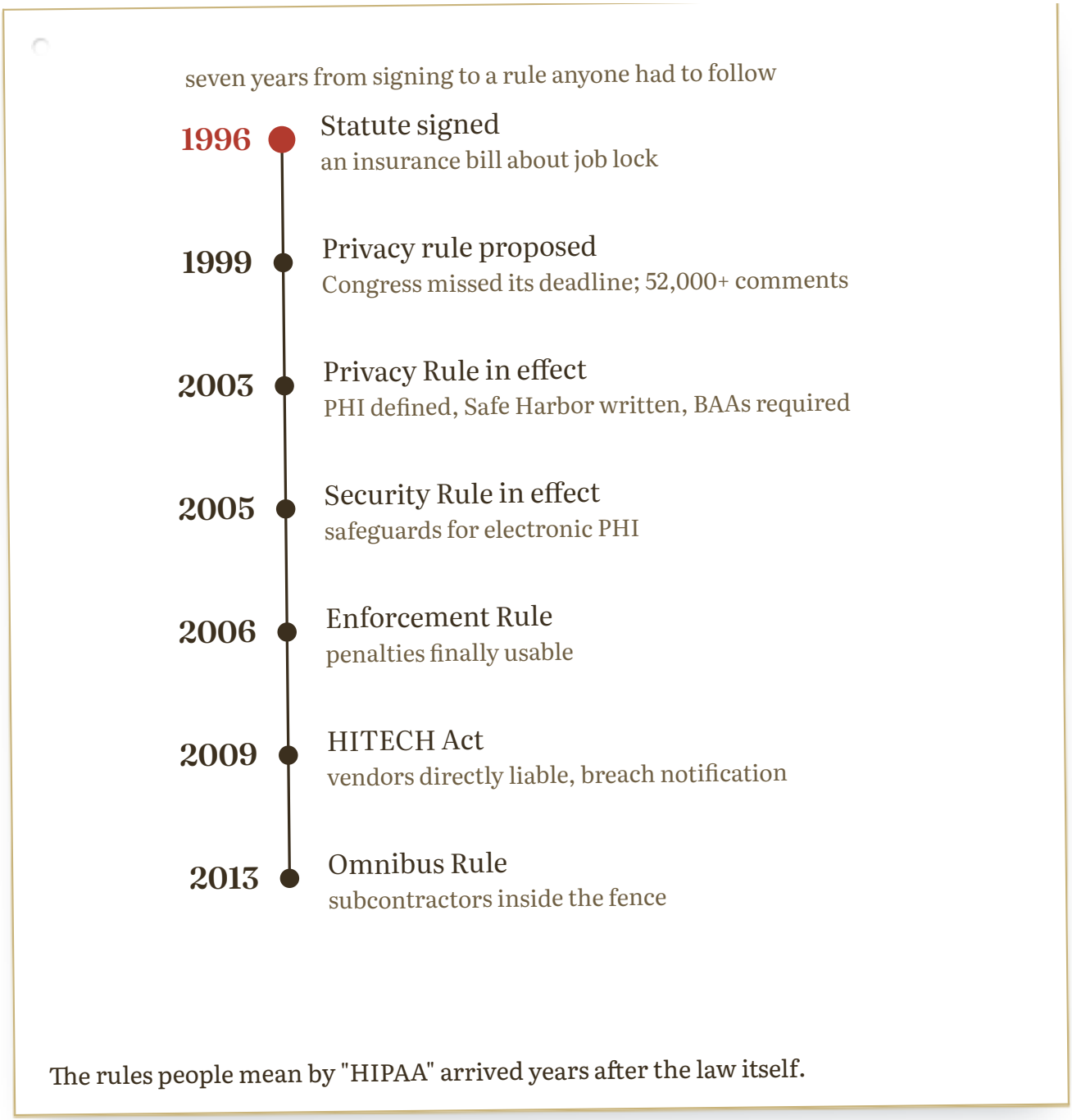
> **Why it matters:** if your product receives data from a hospital, there is a good chance it arrives as HL7 v2 messages.

HIPAA

Health Insurance Portability and Accountability Act, 1996

What it is: the US law whose Privacy Rule and Security Rule govern how health data is used and protected.

Where it came from: it started as an insurance law. Its main goal was letting people change jobs without losing coverage. Because that raised insurers' costs, Congress added rules to cut fraud and standardize claims processing. Privacy was a late addition: the law said that if Congress did not pass a health privacy law within three years, HHS had to write regulations itself. Congress did not, so HHS did.



The rules people mean by "HIPAA" arrived years after the law itself.

The Privacy Rule took effect in 2003 and the Security Rule in 2005. Real enforcement started in 2006. Later laws (HITECH in 2009, the Omnibus Rule in 2013) added breach notification and extended liability to vendors.

> **Why it matters:** there is no HIPAA certification. The law never created one. A product can be built to satisfy the rules, but nobody can certify it.

PHI

Protected Health Information, 2000

What it is: any health information that can be tied to a person. Names, dates, record numbers, and the medical content attached to them. ePHI is the same thing in electronic form, which is what the Security Rule covers.

Where it came from: the Privacy Rule needed a defined term for what it protected. The formal phrase is "individually identifiable health information." PHI is the short form everyone uses.

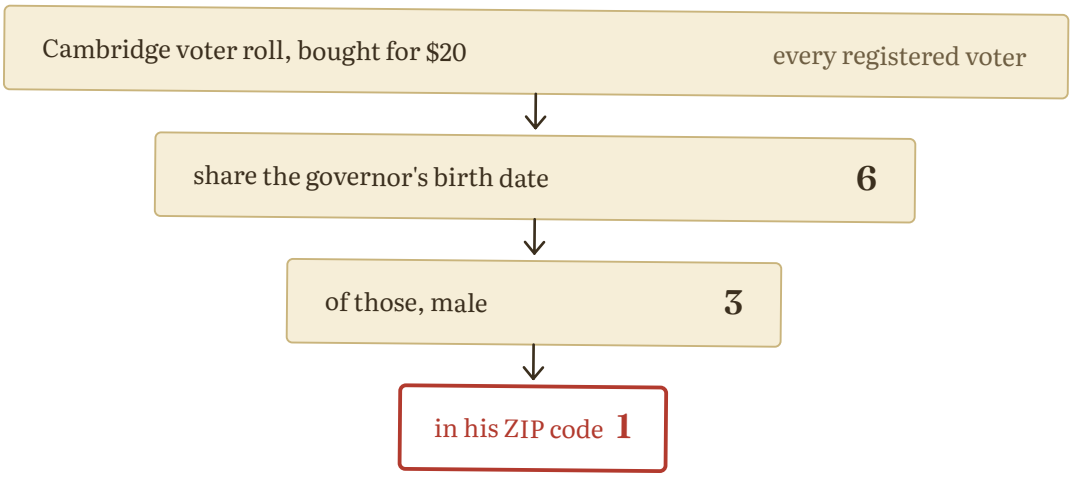
> **Why it matters:** PHI is the thing that must not leak, must not be logged carelessly, and must not reach a model or a vendor without safeguards.

Safe Harbor

HIPAA de-identification method, 2000

What it is: the rule for turning PHI into non-PHI. Remove eighteen types of identifiers and have no reason to believe the rest could identify anyone, and the data is no longer protected.

Where it came from: in 1997, MIT graduate student Latanya Sweeney showed why removing names is not enough. Massachusetts had released hospital records with names and addresses stripped. Sweeney bought the Cambridge voter list for twenty dollars and matched the two datasets on ZIP code, birth date, and sex. Only six voters shared the governor's birthday; three were men; one lived in his ZIP code. She mailed the governor his own medical file.



Three fields, none of them a name, narrowed a whole city to one person.

Her research showed that about 87 percent of Americans can be identified from just those three fields. The Safe Harbor list is a direct response. It removes names, any geography smaller than a state, all date parts except the year, ages over 89, phone numbers, email addresses, Social Security numbers, MRNs, account numbers, device IDs, URLs, IP addresses, biometrics, photos, and any other unique code. ZIP codes may keep their first three digits only where that area has more than 20,000 people.

> **Why it matters:** this is why a de-identifier deletes "March 9" but keeps "daily." A date can identify someone. A dosing frequency cannot.

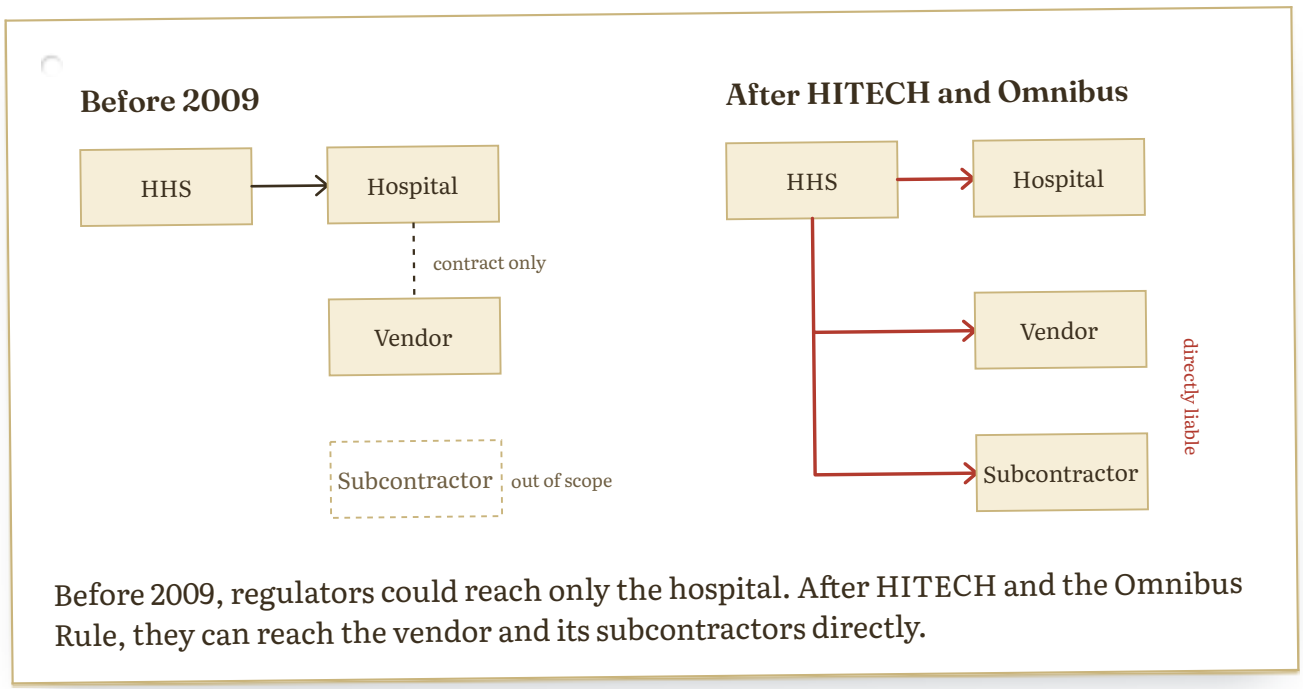
BAA

Business Associate Agreement, 2003

What it is: the contract a hospital must have with any outside company that handles PHI on its behalf. It sets out what the vendor may do with the data and how it must protect it.

Where it came from: the Privacy Rule required it from 2003, but for

the first six years the vendor was only accountable to the hospital, under the contract. Regulators could not go after the vendor directly. The HITECH Act of 2009 changed that, and the 2013 Omnibus Rule extended it to subcontractors: your hosting provider and your AI supplier are covered too.



Before 2009, regulators could reach only the hospital. After HITECH and the Omnibus Rule, they can reach the vendor and its subcontractors directly.

> **Why it matters:** you are a business associate the moment you receive PHI, signed contract or not. Refusing to work without a signed BAA is good practice, but the liability exists either way.

NPI

National Provider Identifier, 2007

What it is: a ten-digit number, one per clinician or healthcare organization, used on every standard transaction.

Where it came from: HIPAA's claims-standardization section required national identifiers, because a standard claim is useless if every insurer identifies the same doctor with a different number. The NPI became mandatory in 2007.

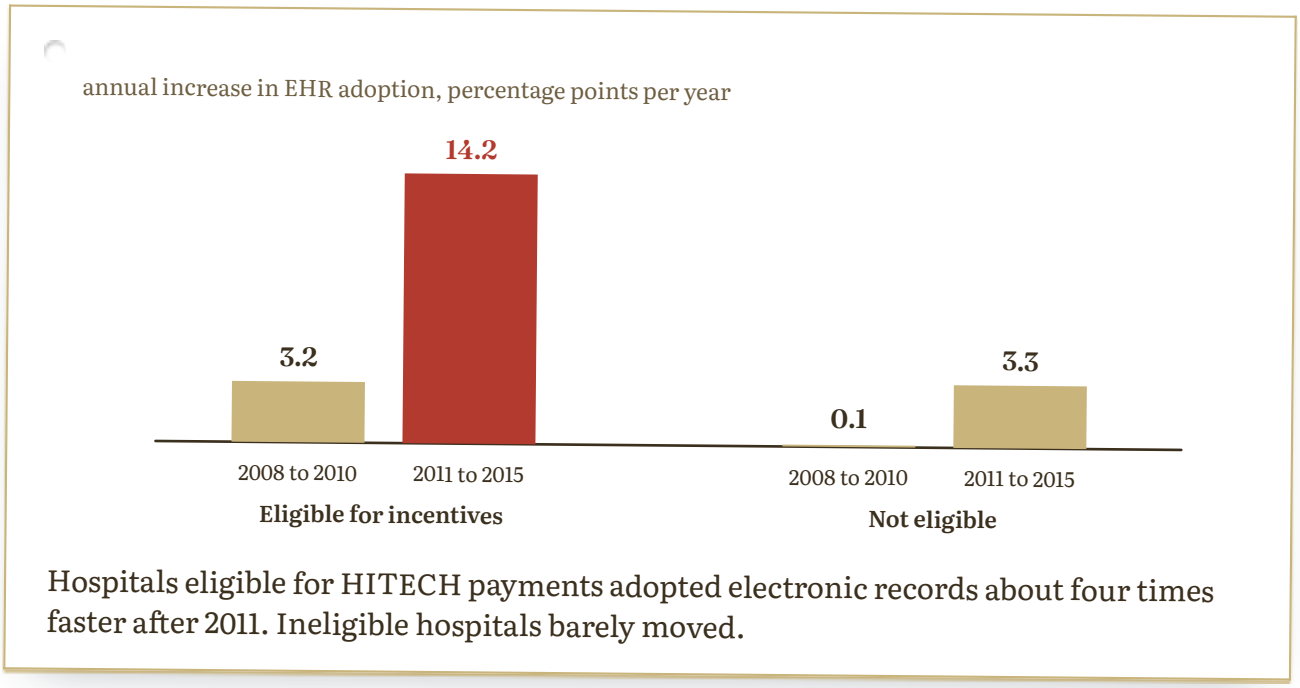
> **Why it matters:** NPIs are public and searchable in an open registry. MRNs are protected. Same kind of key, opposite privacy rules.

HITECH

Health Information Technology for Economic and Clinical Health Act, 2009

What it is: the law that paid hospitals to adopt electronic records and made vendors directly liable for breaches.

Where it came from: by 2008, most US hospitals were still on paper. HITECH, passed inside the 2009 stimulus package, offered about $27 billion in Medicare and Medicaid payments to hospitals that adopted a certified EHR and could show they were using it. A 2017 study compared hospitals that qualified for the money with those that did not.



Hospitals eligible for HITECH payments adopted electronic records about four times faster after 2011. Ineligible hospitals barely moved.

> **Why it matters:** HITECH is why there is electronic data for your product to read. It is also the law that made you liable for what you read.

FHIR

Fast Healthcare Interoperability Resources, 2011

What it is: a web API standard for health records. A patient, a lab result, or a medication is a JSON object at a predictable URL, fetched over ordinary HTTP.

Where it came from: HL7 version 3 had failed to ship after a decade. In 2011 HL7 asked for a fresh start, and Grahame Grieve, a programmer in Melbourne, wrote a proposal over one summer using plain web conventions. HL7 adopted it, renamed it FHIR (pronounced "fire"), and Grieve insisted it be open source.

HL7 v2, 1987

```
MSH|^~\&|LAB|HOSP|...|ADT^A01|
PID|1||000001^^^HOSP^MR||
DOE^JANE||19800312|F|||
PV1|1|I|WARD^12^1|||...
```

pipes and carets, position is meaning

FHIR, 2011

```
GET /Patient/000001

{ "resourceType": "Patient",
"identifier": [{"value": "000001"}],
"birthDate": "1980-03-12",
"gender": "female" }
```

a URL and a JSON object a web developer can read

The same patient as an HL7 v2 message and as a FHIR resource. Both contain the MRN, and both are PHI.

> **Why it matters:** FHIR is how modern apps read and write to an EHR, and it is the format US interoperability rules now require.

ICD-10-CM

ICD, tenth revision, Clinical Modification, 2015

What it is: the current US diagnosis code set. "CM" means the US-specific version, which is far more detailed than the WHO base version.

Where it came from: the WHO released ICD-10 in the 1990s. Australia and Canada moved in 1998 and 2000. The US kept ICD-9-CM for clinical use until October 1, 2015, after Congress delayed the switch several times.

> **Why it matters:** any US data from before late 2015 uses ICD-9 codes. Mixed datasets need a mapping.

Quick reference

ICD	Diagnosis codes. WHO standard; US uses ICD-10-CM since 2015.
MRN	Hospital's patient ID. Protected; removed when de-identifying.
CPT	Procedure codes. AMA, 1966.
SOAP	Note structure: Subjective, Objective, Assessment, Plan.
EMR / EHR	The digital chart. Terms are interchangeable in practice.
HL7	Hospital messaging standard. Version 2 is everywhere.
HIPAA	The 1996 law. Privacy Rule 2003, Security Rule 2005. No certification exists.
PHI / ePHI	Identifiable health information. What the rules protect.

Safe Harbor	Remove 18 identifier types and data is no longer PHI.
BAA	Required vendor contract. Vendors directly liable since 2009, subcontractors since 2013.
NPI	Public ten-digit provider ID, mandatory since 2007.
HITECH	2009 law that funded EHR adoption and extended liability to vendors.
FHIR	Web API standard for health records. Open source, 2011.
ICD-10-CM	Current US diagnosis codes. Pre-2015 data uses ICD-9-CM.

---

1. WHO, "History of the development of the ICD"; AAPC, "The Rules Are Changing: ICD's Continued Evolution" (2023).
2. Mayo Clinic Proceedings, "Patient Records at Mayo Clinic: Lessons Learned From the First 100 Patients" (2008); Mayo Clinic History & Heritage on Henry Plummer.
3. Weed, "Medical Records That Guide and Teach," New England Journal of Medicine (1968); Elation Health and Patagonia Health histories of the EHR.
4. Wikipedia, "HL7 International"; HL7, "FHIR Celebrates 10th Birthday"; Medblocks, "A History of HL7."
5. HHS, "Summary of the HIPAA Privacy Rule"; HIPAA Journal, "The Comprehensive History of HIPAA" and "HIPAA Compliance for Business Associates."
6. Barth-Jones et al., "Re-Identification Risk in HIPAA De-Identified Datasets" (2018); Harvard Gazette, "You're not so anonymous" (2011).
7. Vedder Price and Crowell & Moring alerts on the 2013 HIPAA Omnibus Final Rule.
8. Adler-Milstein and Jha, "HITECH Act Drove Large Gains in Hospital Electronic Health Record Adoption," Health Affairs (2017).