

# WPROWADZENIE DANE BIOMEDYCZNE

---

*Informatyka w medycynie, 2017-2018*

# Literatura

- E. H. Shortliffe, J.J. Cimino (red.): *Biomedical Informatics: Computer Applications in Health Care and Biomedicine*. Springer, 2014; rozdziały 1, 2, 13.
- E. Hozlinger: *Biomedical Informatics. Discovering Knowledge in Big Data*, Springer, 2014; rozdziały 1, 2.

# INFORMATYKA (BIO)MEDYCZNA

---

# Ewolucja dziedziny

lata 70-te

## 1. *Medical computer science (CS)*

- zastosowanie technik obliczeniowych w medycynie
- nacisk na metody, algorytmy...

lata 80-te

## 2. Informatyka medyczna (*medical informatics*)

- nacisk na informację (stąd nazwa) i jej przetwarzanie
- wyjście poza „czyste” metody obliczeniowe

*Informatics = information  
+ automatics*

lata 90-te

## 3. Bioinformatyka (*bioinformatics*)

- przetwarzanie informacji związanej z biologią i genetyką (→ Human Genome Project)

XXI wiek

## 4. Informatyka biomedyczna

- połączenie informatyki medycznej i bioinformatyki w jedną dziedzinę
- przetwarzanie informacji medycznej oraz biologicznej

# Definicja dziedziny

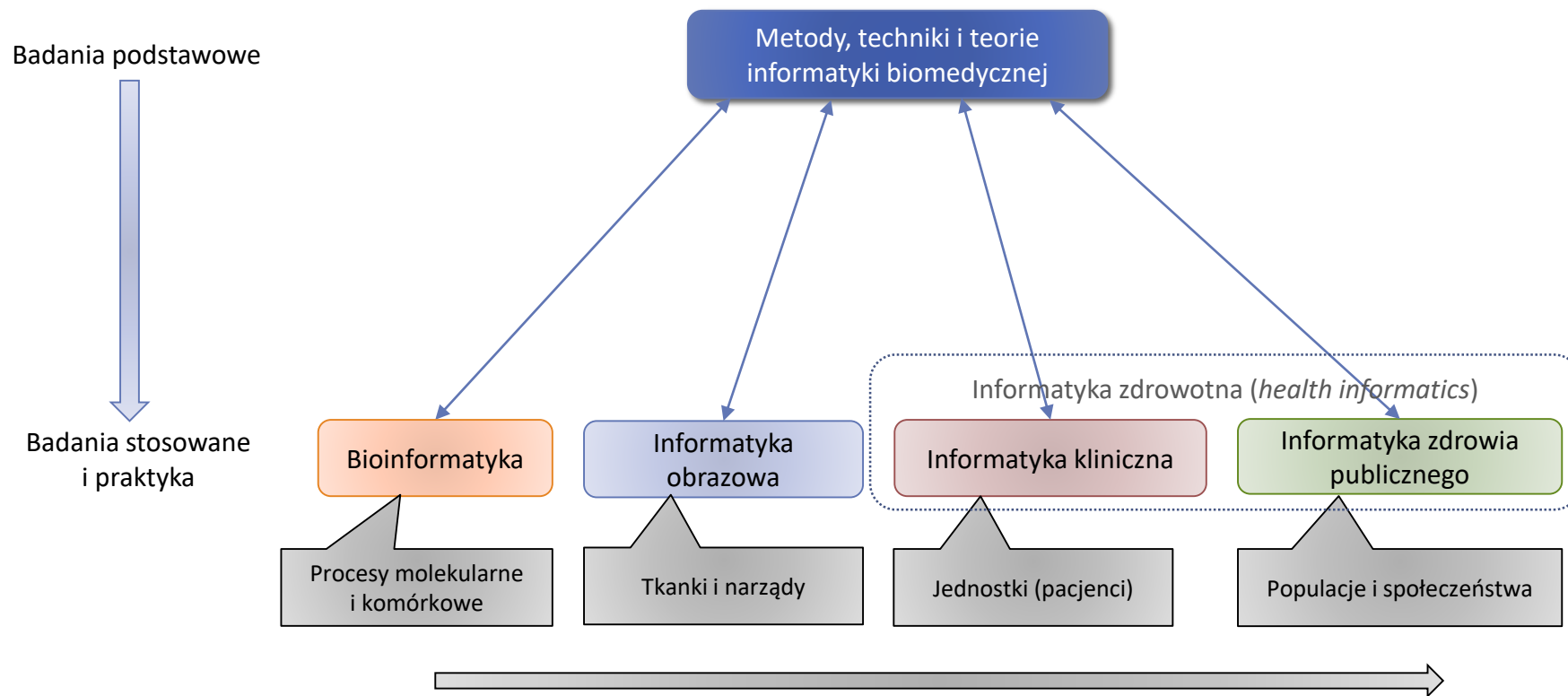
Oficjalna definicja AMIA (American Medical Informatics Association)

**Informatyka biomedyczna** jest dziedziną **interdyscyplinarną** zajmującą się badaniem i efektywnym wykorzystaniem **biomedycznych danych, informacji wiedzy** w celu prowadzenia badań naukowych i rozwiązywania problemów (**podjmowania decyzji**), której celem nadrzędnym jest **poprawa stanu zdrowia**

1. *Zakres*: od molekuł i komórek poprzez jednostki (pacjentów) aż do społeczeństw i populacji
2. *Teorie i metodologie*: tworzenie, przechowywanie, wykorzystanie i współdzielenie biomedycznych danych, informacji i wiedzy
3. *Technologie*: informatyka, telekomunikacja, ...
4. *Kontekst społeczny*: ludzie jako dawcy danych i odbiory wyników (kwestie etyczne)

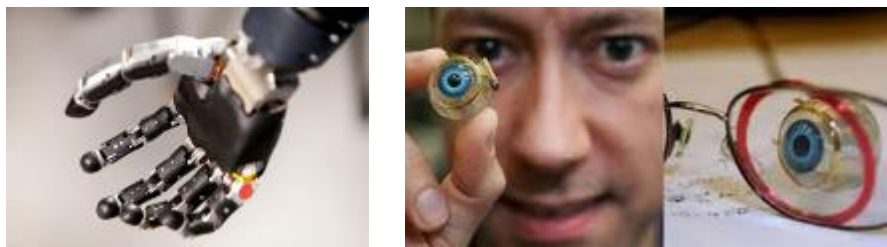
**Informatyka biomedyczna** = *biomedical CS* + analiza i wspomaganie decyzji + statystyka + kognitywistyka + informatologia + zarządzanie (badania operacyjne)

# Badania podstawowe i zastosowania



# Inżynieria biomedyczna

- Starsza niż informatyka biomedyczna – lata 70-te (→ podobnie jako *medical CS*)
- Podstawowy cel – budowa urządzeń (aparatów) medycznych, protez, tworzenie zaawansowanych materiałów i tkanek



- Płynna granica między inżynierią a informatyką biomedyczną – zaawansowane urządzenia pomiarowe z rozbudowaną „logiką”

# Krótką historia

- 1920...
  - systemy do przetwarzania kart dziurkowanych
  - badania epidemiologiczne, monitorowania populacji
- 1970...
  - gromadzenie i przechowywanie danych
  - pierwsze systemy szpitalne i wspomaganie decyzji
- 1985...
  - systemy rozproszone (*health care networks*), początki telemedycyny i telematyki
- 1995... **Web era**
  - aplikacje i usługi webowe
- 2005... **Ambient era**
  - systemy wszechobecne i otaczające (*pervasive & ambient*)
- 2010... **Quality era**
  - kontrola jakości informacji (oraz opieki), *participatory & personalized medicine*





# Personalized Medicine

## Personalized medicine: tailored treatments

Medicine of the present: one treatment fits all    Medicine of the future: more personalized diagnostics

<https://www.cancercommons.org/>

Cancer Commons

Contact Us    Donate

About Us    Patients    Physicians    Knowledge Blog    News

Cancer patients with e.g. colon cancer

Therapy

Effect

**BORN FROM NECESSITY**

Founder and melanoma survivor Marty Tenenbaum is dedicated to helping every patient access the knowledge they need. [Read Marty's story.](#)

**Changing the way the world treats cancer, starting with you.**

Different people respond to a group of patients with different effects (left).

The reason: the genetic makeup and metabolic profile of each individual patient influences the effect of a drug. Personalized medicine takes these individual patterns of cellular and metabolic products into account in the diagnostic phase: **biomarker diagnostics** separates patients into groups with similar characteristics, and provides information on the best individual treatment. This should enable all patients to benefit from their own, "personal" therapy.

**Cancer Commons** is a patient-centric not-for-profit network of patients, physicians, and scientists that help identify the best options for treating an individual's cancer. Members contribute to and benefit from a uniquely dynamic knowledge base of case histories, treatment insights and rationales, clinical trials, and patient outcomes and experiences.

Each case builds on this collective knowledge, enabling more accurate personalized options for the next community member. By learning from each and every patient in this way, we can optimize and expand the use of available treatments far more efficiently than clinical trials, and rapidly improve outcomes for patients with advanced cancer.

<http://pharma.bayer.com/en/innovation-partnering/research-focus/oncology/personalized-medicine/>

# Główne wyzwania

Nonetheless, the enormous technological advances of the last three decades [...] have all combined to make the routine use of computers by all health workers and biomedical scientists inevitable. **A new world is already with us, but its greatest influence is yet to come.**

Sittig, JAMIA (1994):

1. Zunifikowana terminologia medyczna
2. Kompletny i „powszechny” elektroniczny rekord pacjenta
3. Automatyczne kodowanie informacji tekstowej
4. Automatyczna analiza danych pacjentów  
(→ charakterystyka chorób i sposobów ich leczenia)
5. Ujednolicony, intuicyjny i „przewidujący” interfejs użytkownika
6.  Human Genome Project
7.  Kompletna trójwymiarowa reprezentacja ludzkiego ciała
8. Integracja nowych podejść do przetwarzania informacji w organizacjach
9. Kompletny/wszechstronne systemy wspomaganie decyzji



# Główne wyzwania

Nonetheless, the enormous technological advances of the last three decades [...] have all combined to make the routine use of computers by all health workers and biomedical scientists inevitable. **A new world is already with us, but its greatest influence is yet to come.**

20+ lat później:

1. Szybsze przejście od teorii do praktyki (obecnie ok. 17 lat)
2. Zastosowanie metody integracji i fuzji danych w praktyce klinicznej
3. Powiązanie i równowaga między standaryzacją i personalizacją
4. Uniwersalny, adaptacyjny i adaptowalny interfejs użytkownika
5. Interaktywne metody odkrywania wiedzy (zwłaszcza w przypadku danych nieustrukturalizowanych)
6. Systemy mobilne i wszechobecne, dostępne „przy pacjencie”

# DANE, INFORMACJA I WIEDZA

---

# Dane biomedyczne

- Dana (*datum*) to pojedyncza obserwacja opisana następującymi elementami
  1. **pacjent**, którego dotyczy
  2. **parametr**, który jest obserwowany/mierzony (np. waga, ciśnienie krwi, także parametry subiektywne)
  3. **wartość** parametru (np. waga 80 kg, wymiar wątroby 25 cm, ...)
  4. **czas** dokonania pomiaru (precyzja zależna od parametru...)
  5. **metoda** dokonania pomiaru (np. odpowiedź pacjenta, analizator laboratoryjny,...)

# Dodatkowa charakterystyka danych

- Kontekst pozyskania danych
  - Co robił pacjent podczas pomiaru?
  - Czy pomiar był dokonany w spoczynku, czy po wysiłku?
  - Jakie urządzenie pomiarowe wykorzystano?

Uwzględnienie kontekstu może wiązać się ze modyfikacją obserwowanej wartości podczas analizy (kontekst → **modyfikator**)

- Pewność (lub niepewność) związana z pomiarem
  - Czy pacjent jest w stanie precyzyjnie odpowiedzieć na pytanie?
  - Czy lekarz jest w stanie precyzyjnie zinterpretować obserwację?

Poprawa pewności poprzez zbieranie dodatkowych danych (przetarg między „zyskiem” na pewności a kosztami związanymi z pozyskaniem danych)



# Informacja i wiedza biomedyczna

- **Informacja** to zestaw danych, które zostały wybrane i odpowiednio zorganizowane w celu dalszej prezentacji/analizy
- **Wiedza** to uogólnienia/wnioski uzyskane z formalnej lub nieformalnej (→ *rules of thumb*) analizy dostępnej informacji

*Dane:* (1) pomiary ciśnienia dla X, (2) diagnoza zawału serca dla X



*Informacja:* zebrane i połączone dane (1) i (2) dla pacjenta X



*Wiedza:* pacjenci z nadciśnieniem mają większe ryzyko zawału serca

*Wiedza:* dziecko skaczące po łóżku nie ma zapalenia wyrostka pomimo objawów 😊



# Dane a informacja

- Płynna granica między danymi a informacją, zależna od sposobu wykorzystania

Jeden pomiar ciśnienia krwi → **dana**

Zestaw danych z pewnego okresu wykazujący przekroczenie norm → **informacja** o nadciśnieniu

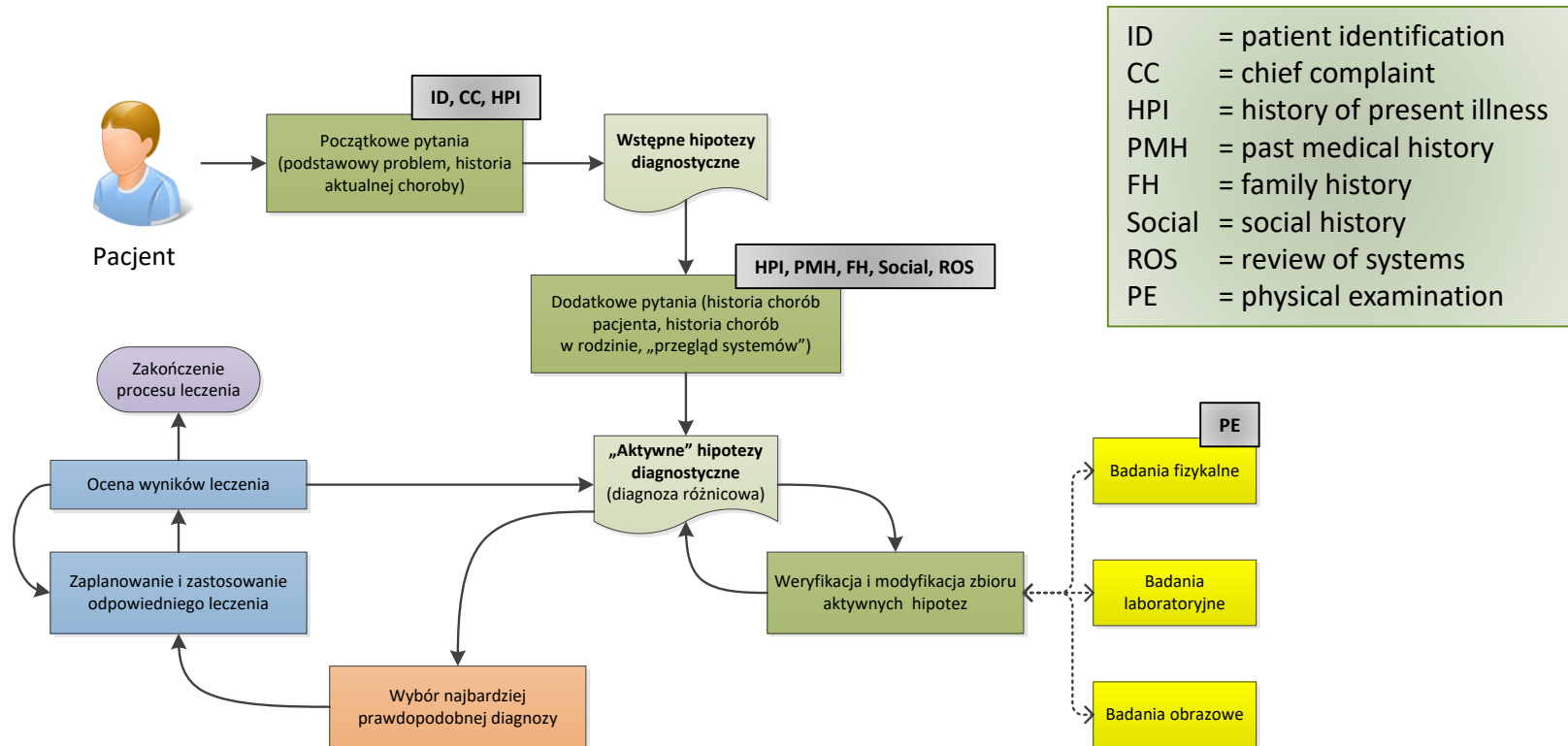
Z punktu widzenia narzędzi wspomaganie decyzji (komputerowych lub „tradycyjnych”) informacja o nadciśnieniu staje się *daną wejściową*

- Rekord/karta pacjenta (→ elektroniczny) to *baza danych* – przechowuje dane i może dostarczać informacji
- Wyniki analizy informacji przechowywane w *bazie wiedzy*

Duże zintegrowane systemy zawierają zarówno bazę danych, jak i bazę wiedzy (→ wykorzystanie mechanizmu wspomaganie decyzji)

# Postępowanie z pacjentem

**Proces hipotetyczno-dedukcyjny** (ang. *hypothetico-deductive*) – iteracyjne stawianie hipotez i weryfikacja ich prawdziwości na podstawie wyników badań



# Dane i hipotezy

Jaki jest związek między obserwacją (daną) a hipotezą, którą „uaktywia”?

- Charakterystyki obserwacji
  - **Czułość** (*sensitivity*) – prawdopodobieństwo tego, że obserwacja występuje w przypadku danego problemu (ważna, ale nie wystarczająca)
  - **Swoistość** (*specificity*) – prawdopodobieństwo tego, że obserwacja **nie** występuje w przypadku **braku** danego problemu

Wysoka czułość → dobry indykator, wysoka swoistość → dobry predyktor

Obserwacja (test) **patognomoniczna** (*patognomonic*) jednoznacznie wskazuje dany problem (swoistość = 100%). Niestety, bardzo rzadka w praktyce klinicznej.

# Dane i hipotezy

Czy swoistość (oraz czułość) wystarczą do postawienia poprawnej hipotezy?

- W procesie stawiania hipotez należy uwzględnić także częstotliwość występowania problemu (*prevalence*)
- **Wartość predykcyjna** (*predictive value*) – zaktualizowane prawdopodob. wystąpienia problemu po dokonaniu obserwacji

$$PV^+ = \frac{\textit{sensitivity} \times \textit{prevalence}}{\textit{sensitivity} \times \textit{prevalence} + (1 - \textit{specificity}) \times (1 - \textit{prevalence})}$$

*sensitivity* = 0.6  
*specificity* = 0.7  
*prevalence*<sub>A</sub> = 0.6  
*prevalence*<sub>B</sub> = 0.3



$PV_A^+$  = 0.75  
 $PV_B^+$  = 0.46

# REKORD PACJENTA

---

# Wykorzystanie rekordu pacjenta

1. Wsparcie dla procesu leczenia → postawienie diagnozy, określenie terapii, weryfikacja jej wyniku
2. Zapewnienie skoordynowanej opieki i zapewnienie komunikacji pomiędzy specjalistami/zespołami (choroby przewlekłe)
3. Monitorowanie pacjenta w celu przewidywania możliwych problemów (medycyna prewencyjna, *preventive medicine*)
4. Dokumentowanie informacji o standardowych zabiegach i działaniach prewencyjnych (np. szczepienia)
5. Dokumentowanie podjętych działań dla ewentualnego postępowania sądowego
6. Agregacja danych i wykorzystanie ich w badaniach klinicznych (badania prospektywne i retrospektywne)

# Wady dokumentacji papierowej

- Problemy z szybkim dostępem i współdzieleniem oraz z czytelnością i kompletnością zawartości (np. brakujące wyniki)
- Nadmiarowość (np. powtarzające się wpisy) i nieefektywna prezentacja danych (mało treści, dużo papieru)
- Brak możliwości automatycznego wyszukiwania i przetwarzania (np. konieczność ręcznego przeglądania podczas agregacji)
- Pasywny charakter papierowego rekordu – brak możliwości kontroli jakości danych oraz monitorowania stanu pacjenta



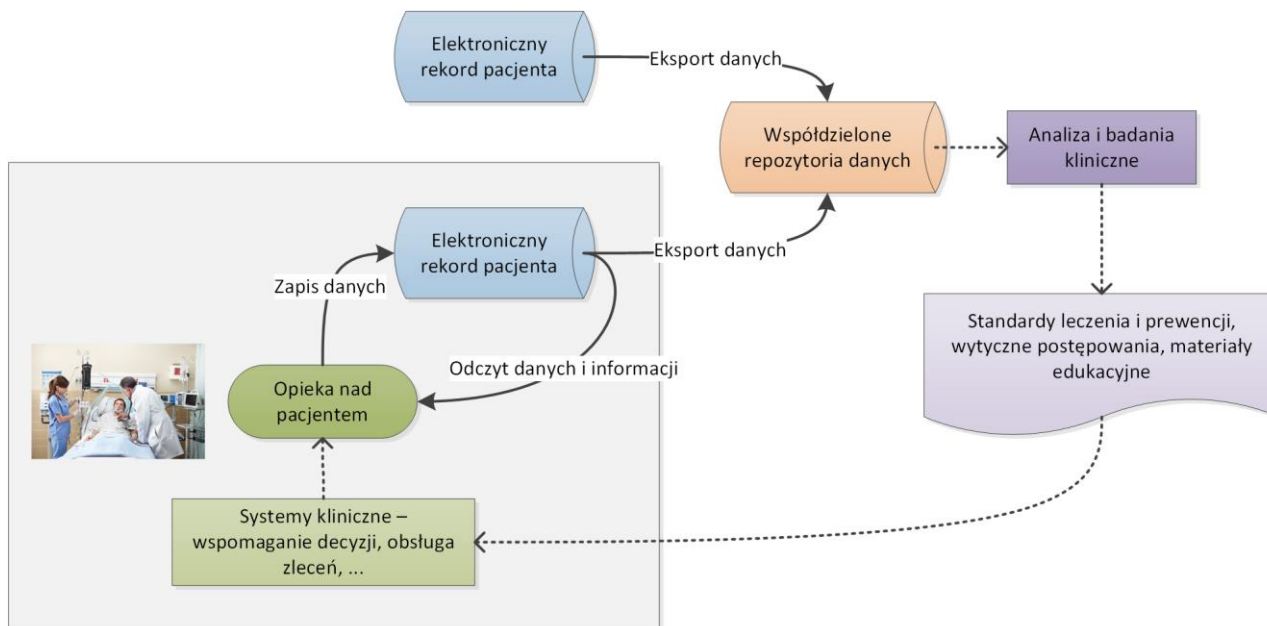
Bardzo ograniczone możliwości wykorzystania danych!



# Dokumentacja elektroniczna

- Dla pacjenta – poprawa jakości opieki i jej bezpieczeństwa
- Dla organizacji – poprawa efektywności i jakości działania

Pierwszy (i istotny) krok w kierunku **uczącego się systemu ochrony zdrowia**  
(*learning health care system*)





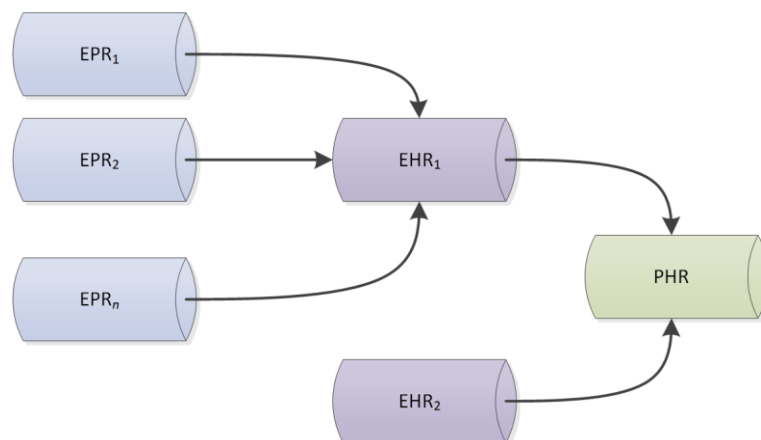
# Powszechny dostęp do danych

- Infrastruktura (*health information infrastructure, HII*) zapewniająca dostęp do **kompletnych** danych zgromadzonych w różnych systemach używanych przez wiele organizacji

**EPR/EMR** (*electronic patient/medical record*) = jedna organizacja

**EHR** (*electronic health record*) = wiele organizacji

**PHR** (personal health record) = wiele źródeł, zarządzanie przez pacjenta



 HealthVault

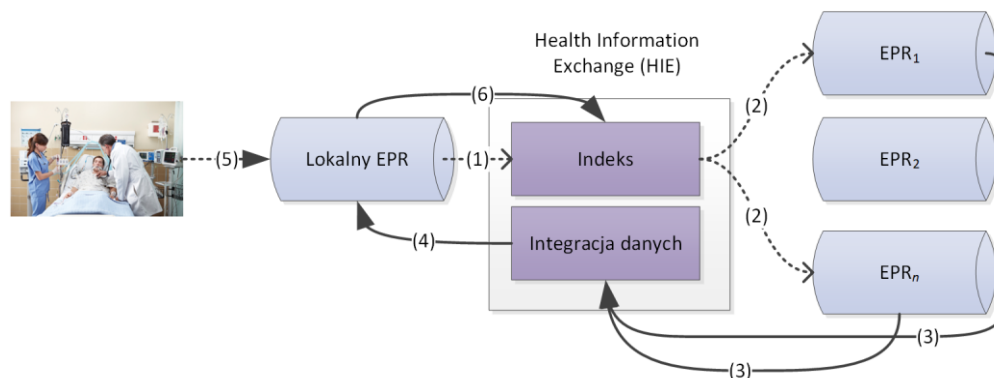
# Zalety HII

- Redukcja liczby błędów popełnianych przez personel medyczny (np. redukcja błędnie przypisanych leków o 55%)
- Wspomaganie testów klinicznych (np. identyfikacja potencjalnych uczestników, automatyczne zbieranie danych)
- Bardziej dokładny monitoring populacji (np. wykrywanie ataków bioterrorystycznych albo wybuchów epidemii)
- Redukcja kosztów opieki zdrowotnej

Zalety podobne jak w przypadku elektronicznego rekordu pacjenta (EPR),  
ale w skali makro!

# Architektury HII

- Architektura zorientowana na instytucje (*institution-centric*) – centralny indeks i rozproszone dane

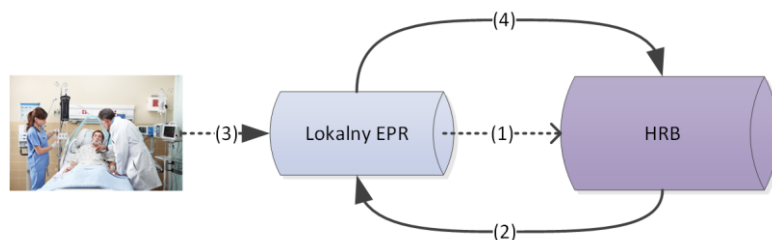


- 👍 Dane pod kontrolą poszczególnych jednostek
- 👎 Utrudnione wyszukiwanie w przypadku złożonych warunków
- 👎 Konieczność monitorowania i utrzymywania niezawodności wielu systemów EPR

# Architektury HII

- Architektura zorientowana na pacjenta (*patient-centric*) → **health record bank (HRB)**

„an independent organization that provides a secure electronic repository for storing and maintaining an individual's life- time health and medical records from multiple sources and assures that the individual always has complete control over who accesses their information”



- 👍 Dane zebrane w jednym miejscu – ułatwione wyszukiwanie
- 👍 Pełna kontrola dostępu do danych przez (świadomego) pacjenta
- 👎 Rozbudowana infrastruktura do przechowywania danych
- 👎 *Single point of failure*

# Wykorzystanie technologii *blockchain*

- Przykładowe rozwiązanie: system MedRec wykorzystujący infrastrukturę Ethereum
- *Blockchain* do zapewnienia dostępu do danych oraz zapisu zmian uprawnień w dostępie do danych (*smart contracts*)
- Dane przechowywane w systemach poszczególnych organizacji (→ *institution-centric*), bloki zawierają wskaźniki do danych
- Nagrodą za generowanie wartości haszujących jest dostęp do zanonimizowanych danych

# MedRec – przykład działania

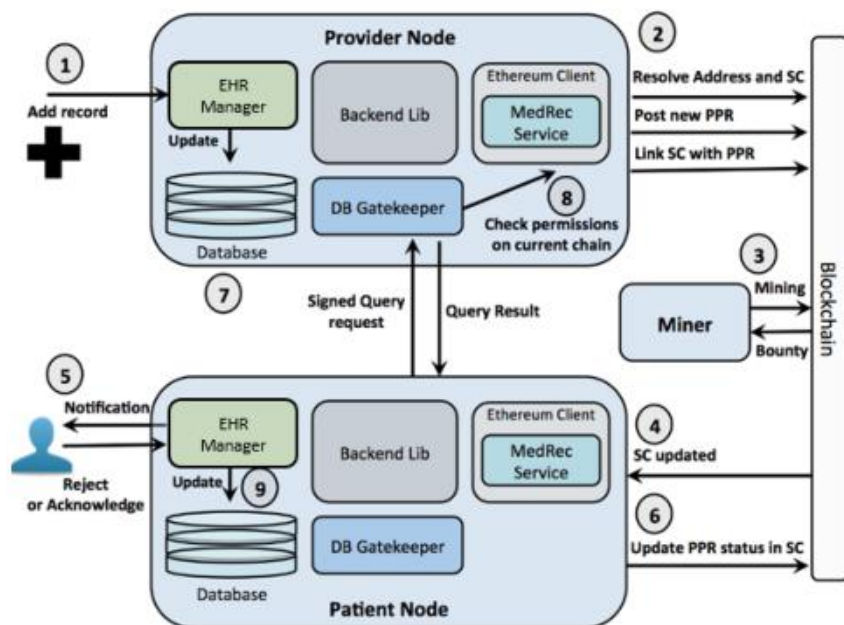
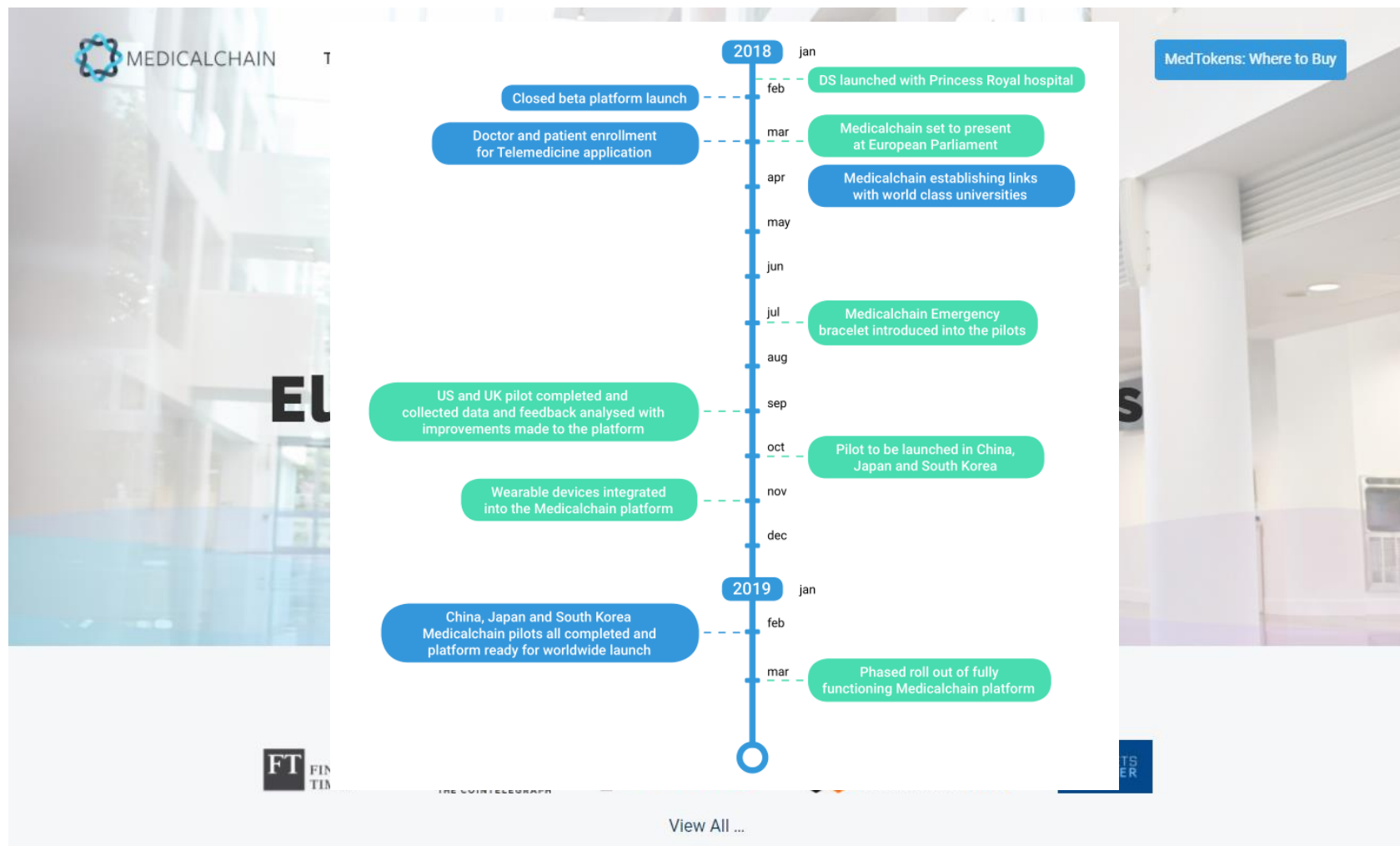


Figure 2. System orchestration example: provider adds a record for new patient.

# Medicalchain

<https://medicalchain.com/en/>



View All ...

# Najważniejsze problemy i wyzwania

1. Zapewnienie odpowiedniego bezpieczeństwa, prywatności oraz anonimowości danych (w przypadku agregacji)
2. Powszechne przyjęcie standardów dla przesyłania i współdzielenia danych (np. HL7)
3. Powszechne przyjęcie standardów modeli danych oraz ich kodowania (np. ICD-10, LOINC, ....)
4. Zapewnienie kontroli jakości agregowanych danych i wykrywanie błędów



# Dobre perspektywy zawodowe

WEDNESDAY, OCTOBER 9, 2013

## Further Evidence That Health IT Job Growth Has Been Underestimated, and Some Ramifications

New data from an analysis of online job postings confirms that employment growth in health information technology (HIT) has even further exceeded projections, driven by funding from the Health Information Technology for Economic and Clinical Health (HITECH) Act. In addition, two other reports show that the job market for those working with electronic health record (EHR) and related systems continues to be strong for employees and challenging for employers.

In the new analysis, Schwartz and colleagues used a comprehensive database of 84 million online job postings, extracted out those related to HIT, and built a model aiming to determine the influence of HITECH [1]. The authors limited their focus to jobs that would be defined in the realm of clinical informatics. Although this was important for their goal of assessing the influence of HITECH, from the perspective of a biomedical informatics graduate program director like myself, this excluded other important informatics jobs, such as those in imaging informatics, bioinformatics, clinical research informatics, and other areas where a graduates of our program are employed.

WILLIAM HERSH



<http://www.billhersh.info/>

FOLLOW BY EMAIL

Email address...

TOTAL PAGEVIEWS



Business Mgmt ▾ Data Security ▾ Pop Health & Data Analytics ▾ Medical IT ▾ HIT Summits ▾

## Report: Health Informatics Labor Market Lags Behind Demand For Workers

December 12, 2014 by Rajiv Leventhal

[f](#) [in](#) [t](#) [G](#) [+](#) [p](#) | Reprints

The demand for health informatics workers is projected to grow at twice the rate of employment overall, but there is strong evidence that the nation already faces a shortage of qualified workers in this field, according to a new report from the Boston-based Burning Glass Technologies.

Health informatics jobs already remain open longer than the national average, a clear sign that employers struggle to fill these positions, according to an analysis of job postings nationwide included in the report. Once primarily clerical, health informatics roles now require a more diverse skill set to meet the demands of an evolving industry that has been reshaped by electronic record-keeping, a shifting regulatory environment that includes the conversion to ICD-10, and "big data," the report concluded.

# DOKUMENTACJA ELEKTRONICZNA W POLSCE

---

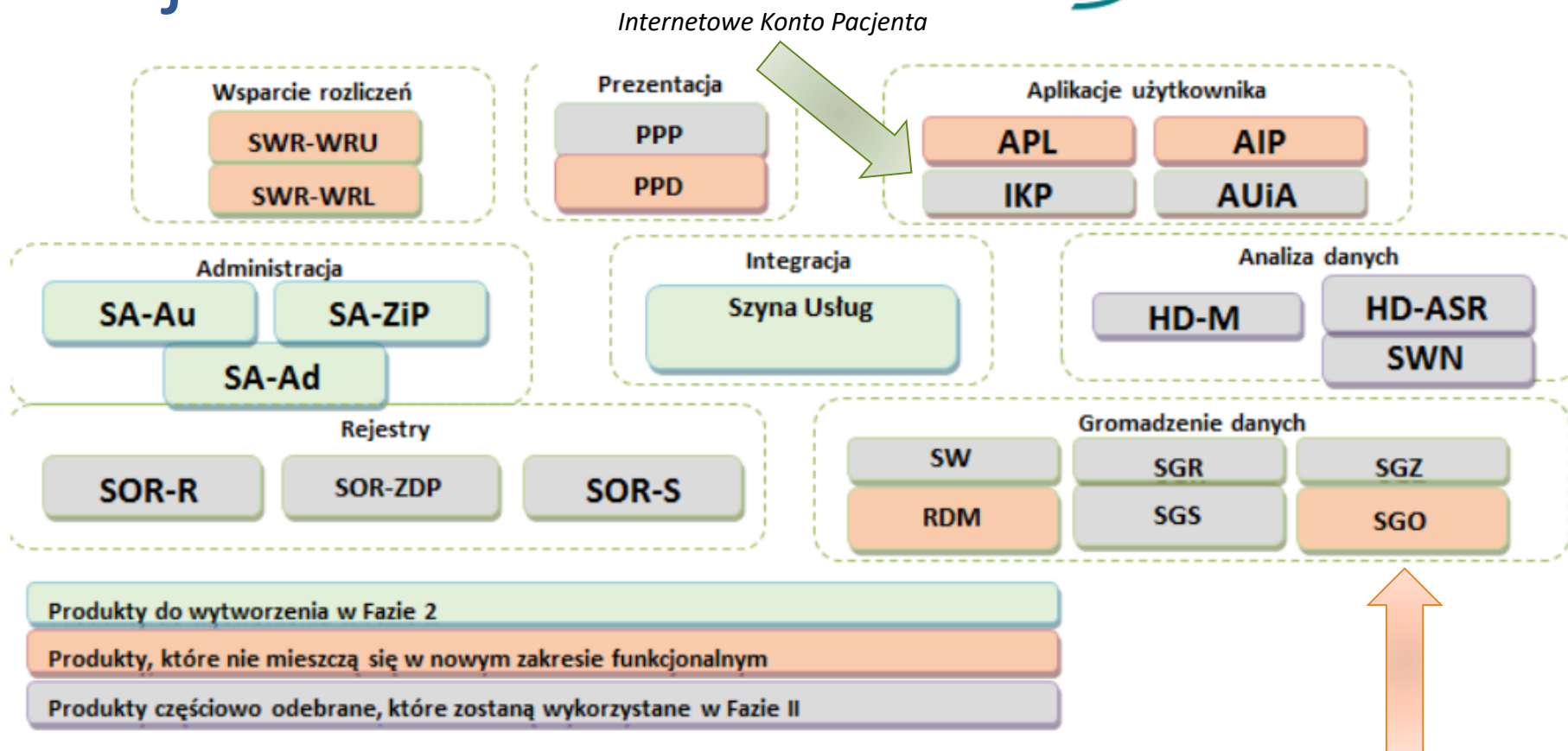
# Praktyczne wdrożenie

- Konieczność prowadzenia dokumentacji elektronicznej od 1.01.2019 (~~1.01.2018~~, ~~1.08.2017~~, ~~1.08.2014~~)
- Wersja elektroniczna dla nowych epizodów/przypadków – dokumentacja papierowa w przypadku kontynuacji leczenia
- Dokumentacja indywidualna, wewnętrzna (np. historia choroby) i zewnętrzna (np. recepta, skierowanie)
- Globalna dokumentacja elektroniczna – HII
  - Projekt P1 (*Elektroniczna Platforma Gromadzenia, Analizy i Udostępniania zasobów cyfrowych o Zdarzeniach Medycznych*) realizowany przez CSIOZ
  - Dokumentacja osobista (HRB) usunięta w fazie II projektu – zamiast tego scentralizowany indeks, kontrola dostępu i usługa udostępniania

# Projekt P1



Centrum Systemów Informatycznych  
Ochrony Zdrowia



*System Gromadzenia Danych Medycznych  
Osobista Dokumentacja Medyczna*

# Zintegrowany Informator Pacjenta

<http://zip.nfz.gov.pl/>

The image shows two overlapping screenshots of the ZIP website. The top screenshot is the home page, and the bottom screenshot shows a detailed view of medical services.

**Home Page (Top Screenshot):**

- Logo: Z I P Zintegrowany Informator Pacjenta
- User: Zalogowany 11251573
- Buttons: Twój portal, Gdzie się leczyć?, Rejestr Usług Medycznych, Prawo do świadczeń
- Navigation: Świadczenia medyczne, Deklaracje POZ, Recepty refundowane, Uzdrawiska, Kolejki oczekujących

**Medical Services Page (Bottom Screenshot):**

**Świadczenia medyczne**

Pokaż opcje wyszukiwania

Data / daty pobytu	Miejsce udzielenia świadczenia
2014-03-18	PRZYCHODNIA LEKARSKA "EUMEDICA" LEK. MED. MICHAŁ DYE GABINET LEKARZA PODSTAWOWEJ OPIEKI ZDROWOTNEJ ul. OBORNICKA 126, 62-002 SUCHY LAS
2013-01-17	PRZYCHODNIA LEKARSKA "EUMEDICA" LEK. MED. MICHAŁ DYE GABINET LEKARZA PODSTAWOWEJ OPIEKI ZDROWOTNEJ ul. OBORNICKA 126, 62-002 SUCHY LAS
2012-04-25	SPECJALISTYCZNE CENTRUM MEDYCZNE OMEGA 045 - PORADNIA DERMATOLOGICZNA ul. ROLNA 17, 62-002 SUCHY LAS

**Recepty refundowane**

Pokaż opcje wyszukiwania

Suma kosztów refundacji wybranych recept: 27,25

Data wystawienia / data realizacji	Miejsce realizacji	Lista leków	Opłata pacjenta	Koszt refundacji	Szczegóły
2012-03-12 2012-03-12	APTEKA "POD LASEM" ul. PLAC GRZYBOWY 7, 62-002 ZŁOTNIKI	SUMAMED	29,07	4,68	<a href="#">pokaż</a>
2011-04-19 2011-04-20	APTEKA "POD LASEM" ul. PLAC GRZYBOWY 7, 62-002 ZŁOTNIKI	SUMAMED	31,97	14,21	<a href="#">pokaż</a>
2009-10-19 2009-10-19	APTEKA "POD LASEM" ul. PLAC GRZYBOWY 7, 62-002 ZŁOTNIKI	AUGMENTIN	16,27	8,36	<a href="#">pokaż</a>

# PODSUMOWANIE

---

# Podsumowanie

- Ewolucja dziedziny:
  - Informatyka medyczna = informatyka + medycyna
  - Informatyka biomedyczna = informatyka medyczna + bioinformatyka
- Dane biomedyczne → informacja → wiedza
- Hipotetyczno-dedukcyjny proces zbierania danych i podejmowania decyzji diagnostycznych i terapeutycznych
- Elektroniczny rekord pacjenta – zalety i wizja uczącego się systemu ochrony zdrowia oraz infrastruktury informacyjnej