

Science For A Better Life

Security vs Freedom – It‘s not a matter of Phylosophy

Dr. Philip Groth

IT Business Partner, R&D IT

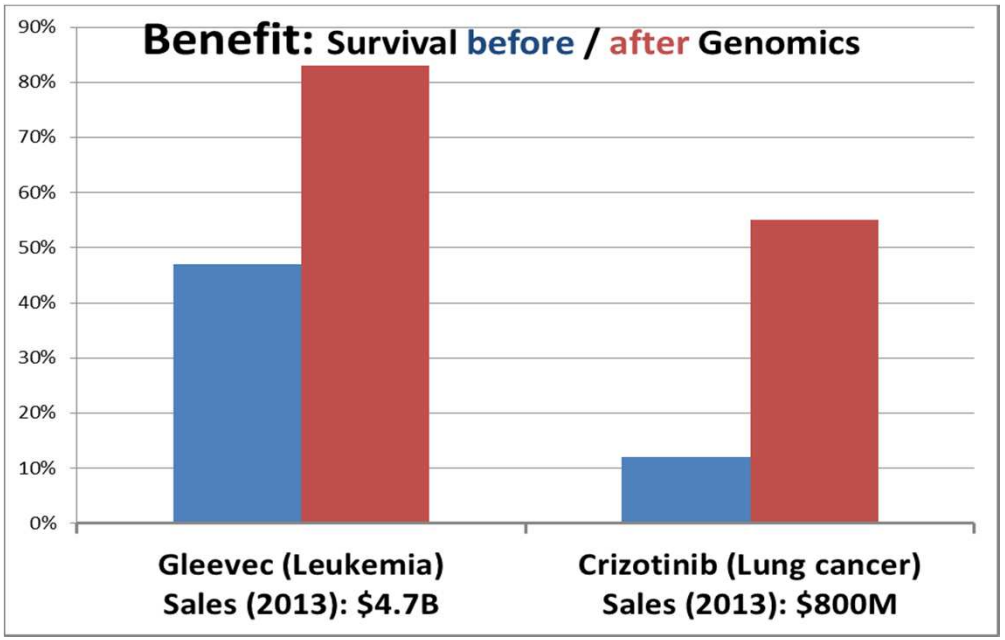
Dr. Nora Manstein

IT Consultant, R&D IT

What is the value of Genomics in Drug Discovery?

Gleevec (1998): BCR-ABL mutated Chronic Myeloid Leukemia

- 5 year survival rate at 89%, with a relapse rate of about 17%
- Before, 30% of patients survived for five years after being diagnosed
- Global sales (2013): $4.7 billion p.a.
- „Gleevec is an exceptional case, and the same success is not likely to be achieved with other cancers any time soon.” (Pray et al., Nat Ed, 2008)



Crizotinib (2010): Xalkori – EML4-ALK mutated Non-small-cell lung cancer

- Before, no survivors within 5 years
- 57% response / 87% disease control rate
- Survival: 1st yr: 74% vs 44%
- Global sales (2013): $800 million p.a.

Sources:
Druker et al., NEJM, 2006.
Kantarjian et al., Blood, 2012.
Shaw et. al., Nat Rev Drug Disc, 2011.
Shaw et al., Lancet Oncology, 2011.

International genome projects generate valuable genomic data

- **ICGC** International Cancer Genome Consortium
  - ICGC joins 51 project teams in 15 countries to study >24K tumor genomes of 50 different (sub-)types.
- **TCGA** The Cancer Genome Atlas
  - To chart the genomic changes involved in more than 20 types of cancer. Received $50M in 2006 from NCI and NHGRI and $275M in 2009 from NIH.
- **PHE** Public Health England
  - Announces to create world's largest cancer database, collating all available data on 350K patients p.a.

University of California Santa Cruz Cancer Genomics Hub

Currently 2,005 TB of TCGA data in 86K files

BBC NEWS HEALTH

12 June 2013 Last updated at 11:06 GMT

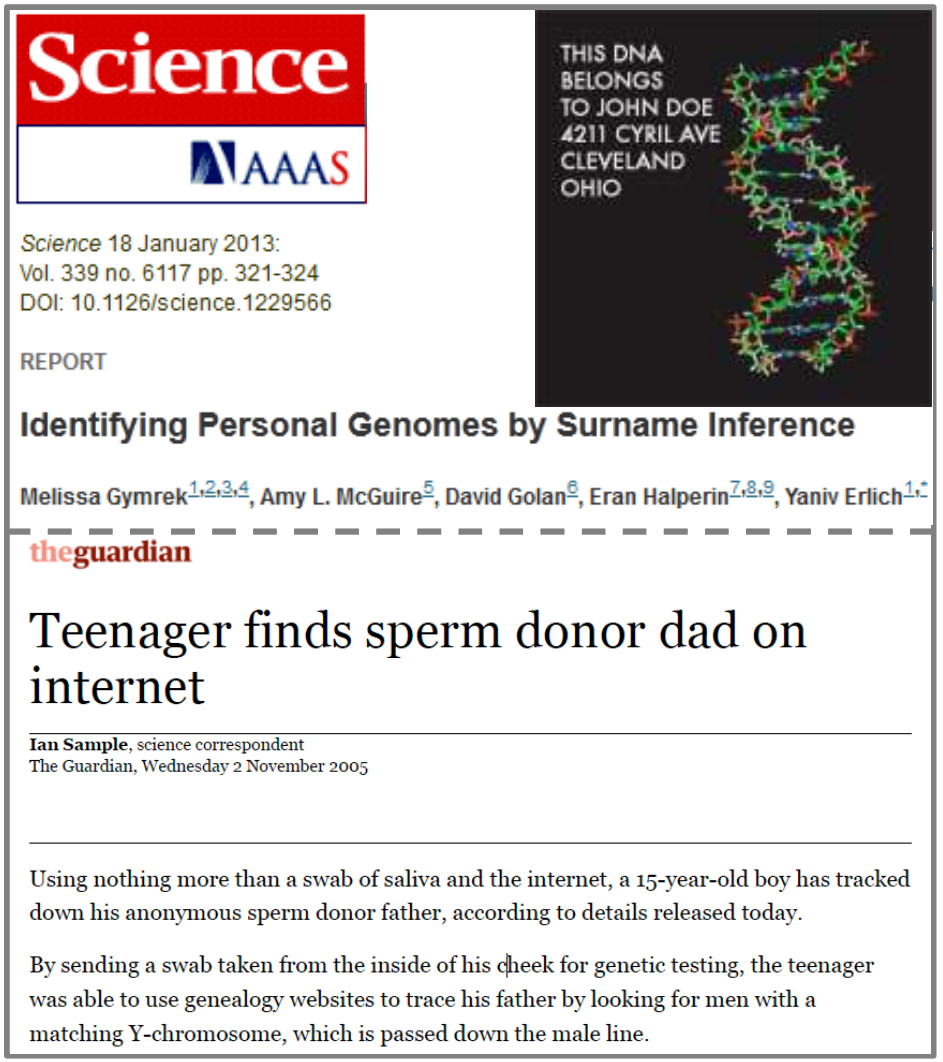
Public Health England to launch largest cancer database

The world's largest database of cancer patients is being set up in England in an attempt to revolutionise care, Public Health England has announced.

It will collate all the available data on each of the 350,000 new tumours detected in the country each year.

Data Privacy needs to be managed

- **Data privacy & security has highest priority**
- Data belonging to a defined person may not be used in contradiction to the person‘s intent;
- Data belonging to a defined person have to be protected from misuse;
- Protection from misuse does always include that noone without a need to access the data gains access;
- Data without individual information are much easier in regard to data protection.



Science

AAAS

Science 18 January 2013:
Vol. 339 no. 6117 pp. 321-324
DOI: 10.1126/science.1229566

REPORT

THIS DNA BELONGS TO JOHN DOE 4211 CYRIL AVE CLEVELAND OHIO

Identifying Personal Genomes by Surname Inference

Melissa Gymrek[1,2,3,4], Amy L. McGuire[5], David Golan[6], Eran Halperin[7,8,9], Yaniv Erlich[1,*]

theguardian

Teenager finds sperm donor dad on internet

Ian Sample, science correspondent
The Guardian, Wednesday 2 November 2005

Using nothing more than a swab of saliva and the internet, a 15-year-old boy has tracked down his anonymous sperm donor father, according to details released today.

By sending a swab taken from the inside of his cheek for genetic testing, the teenager was able to use genealogy websites to trace his father by looking for men with a matching Y-chromosome, which is passed down the male line.

But how?

Bundesdatenschutz-gesetz

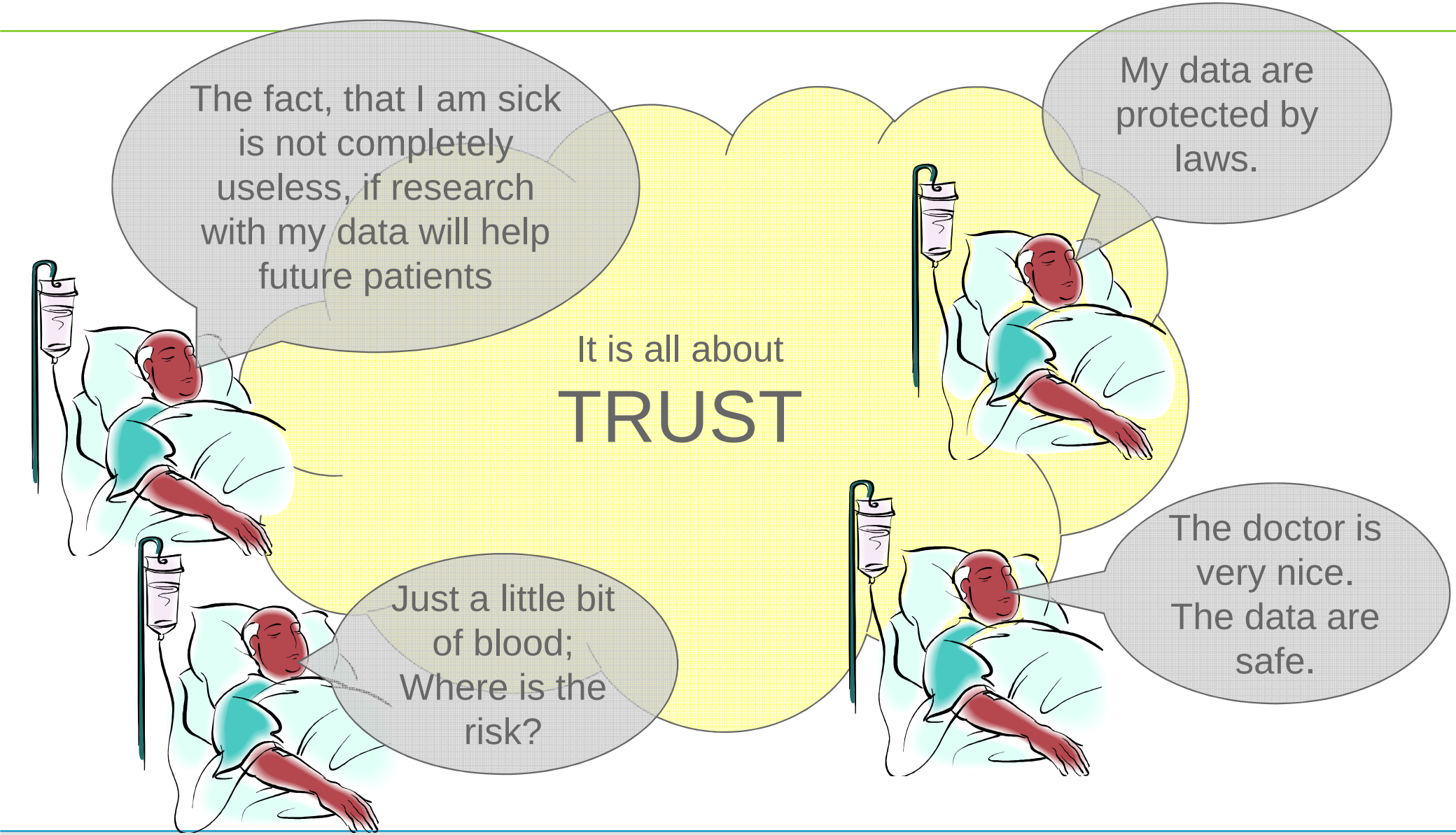
Contracts with third parties

Informed consents

Corporate IT-Security and data privacy rules

YOU!

Need to know the intent of the data owner



Our Goals

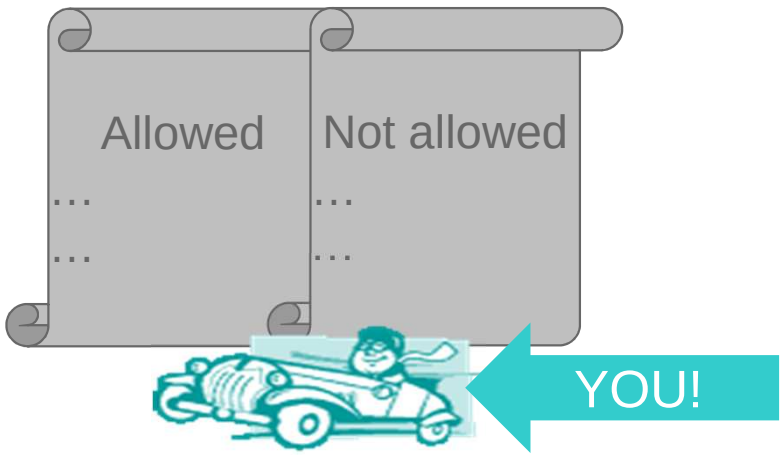
The donator did not provide his data to be protected into being useless.

We must find new products to stay competitive

We must respect the wish of the donator.

We must obey the law

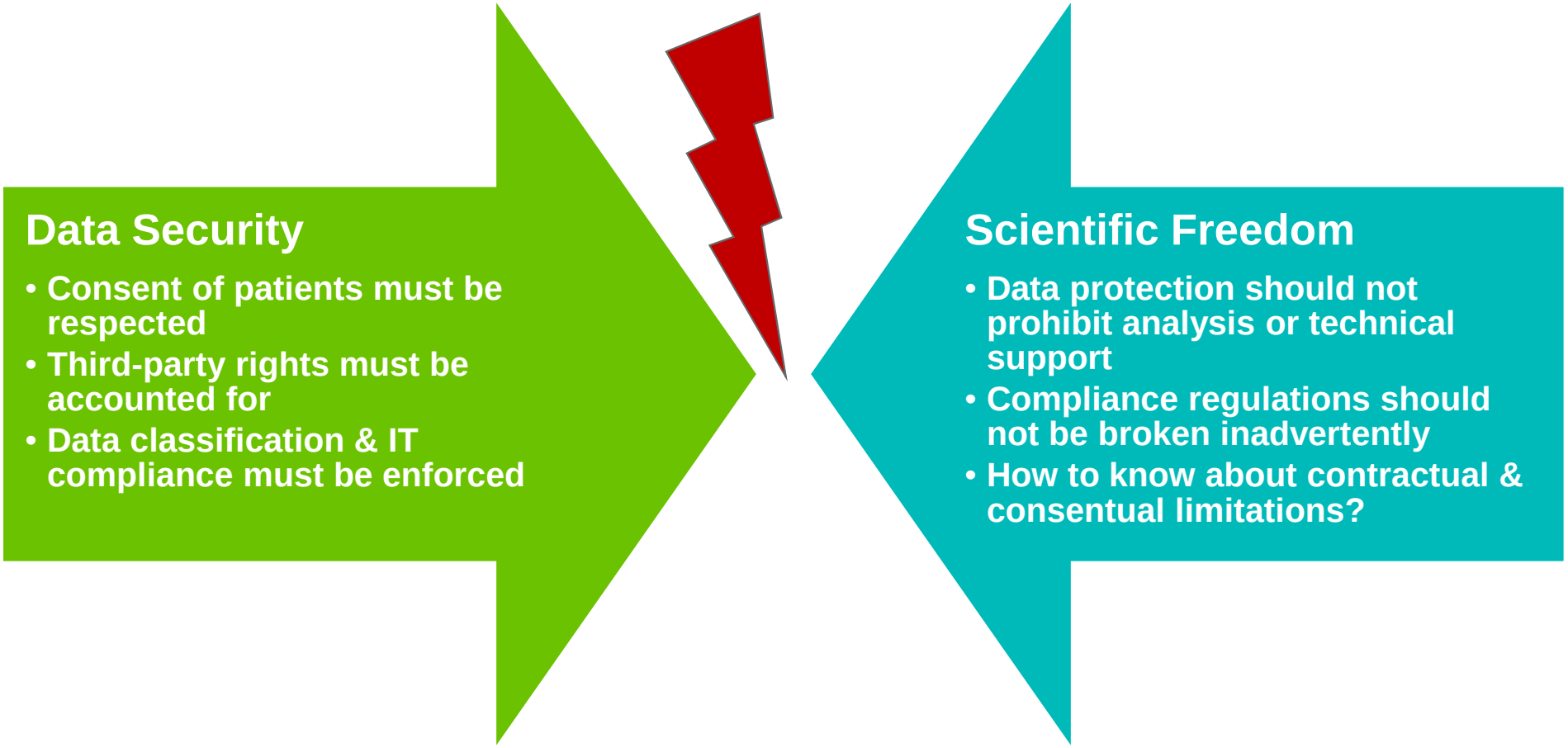
We must honor agreements with research institutes e.g. for publishing data.



Risks in Case of Non-Compliance with Data Privacy Laws

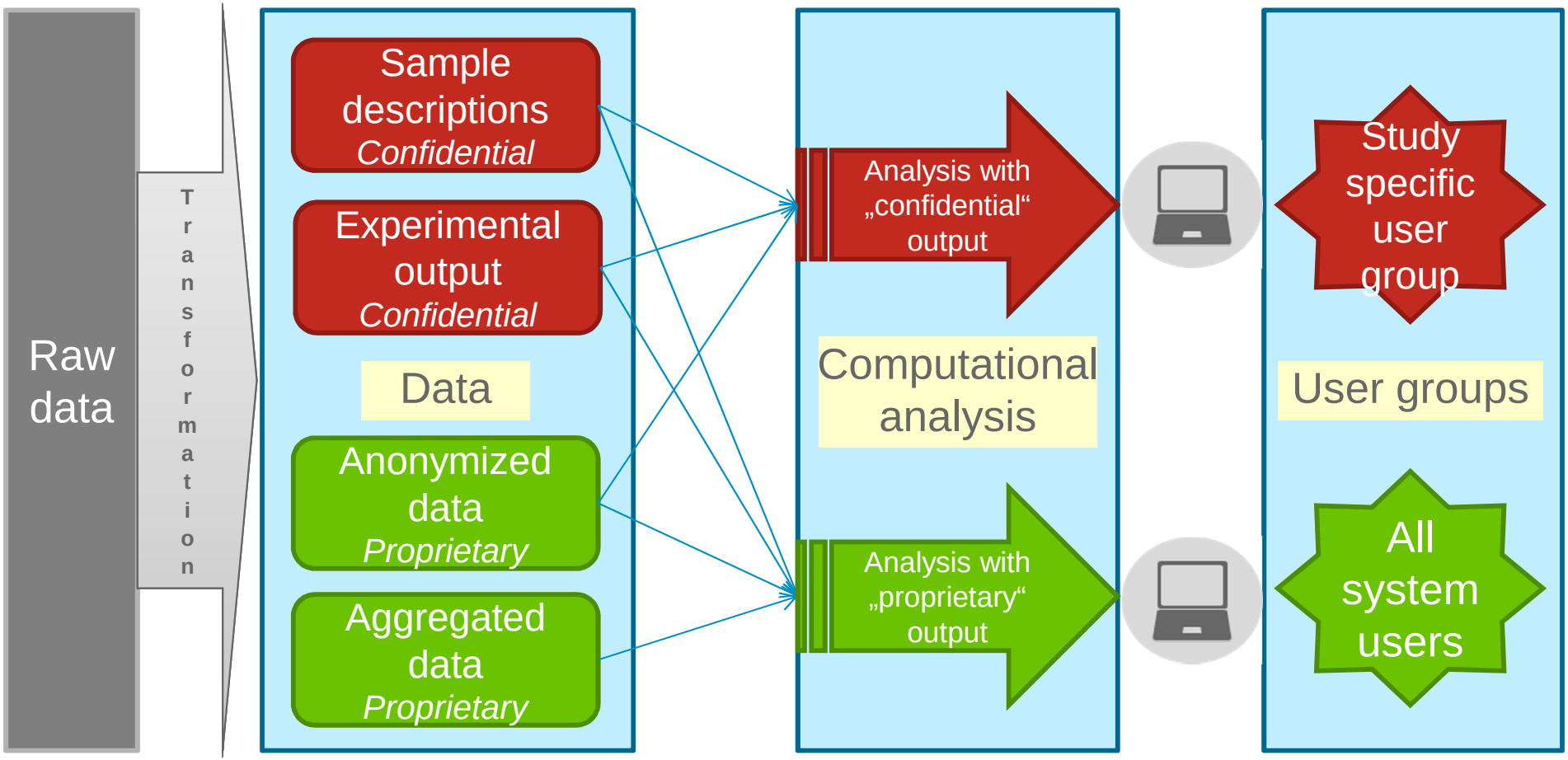
- **Proposed new EU Data Protection Regulation**
  - Fines up to 1M€ or 5% of a company's worldwide annual sales
- **German data protection law**
  - Fines of up to 300k€ per case
  - Imprisonment of up to 2 years in case of wilful misconduct in order to obtain financial benefits
  - Deletion of data/destruction of samples upon administrative act
  - Comprehensive data protection audits by authorities
  - For providers of human samples and data: responsibility under criminal law due to violation of obligation of professional confidentiality/discretion
- **Risk of reputational damages and subsequent strict supervision by authorities**
  - Risk to loose potential partners / sources

Data Security vs Scientific Freedom
How to enable compliant research?



Can R&D IT contribute a solution to this challenge?

What happens to Research Data in Computational Genomics?



- User selects methods and data according to permissions
- „Viewing a single genome“ is also a classified method in this context

Anonymization by Aggregation in Computational Analysis

Confidential

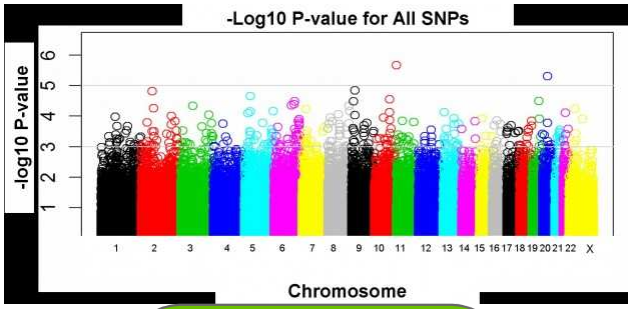
Whole genome variations

combination

Detailed phenotypic descriptions

Population-level summary, e.g. GWAS

Proprietary



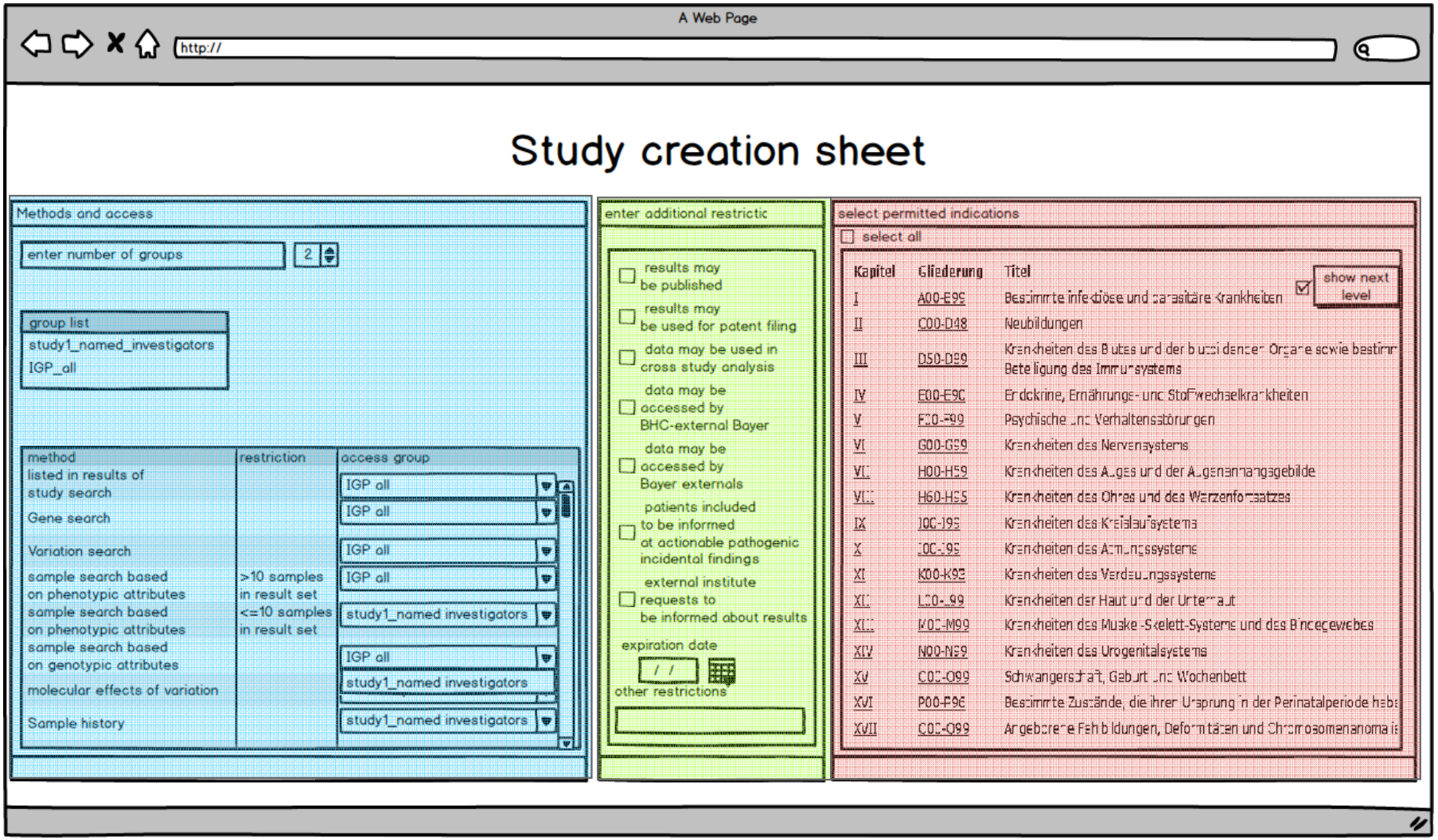
Variations associated with phenotypic properties = aggregated data leaving patient level

Technical Solution for Study Data

User groups with permissions to computational analysis

Studywide usage restrictions

Permitted research indications

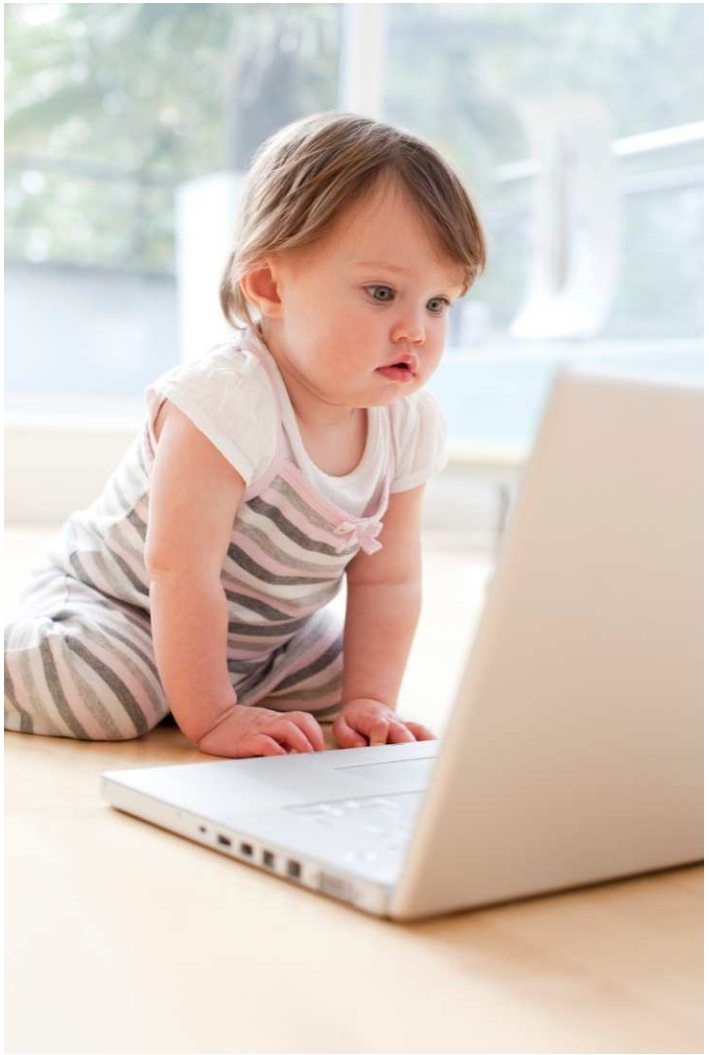


Checks and Balances in Data Submission Process

1. Paperwork: Check, if standard assignments of restrictions fit submitted data 2. Paperwork: Define usage restrictions for data 1. Select permitted research indications 2. Select member of decision body to approve restrictions for data 3. Select, if applicable 1. Expiration date (unlimited as possible value) 2. Results based on data may be published 3. Results based on data may be used for patent filing 4. Use of data in cross study analysis 5. Data may be accessed by employees of the selected organizations 6. Data may be accessed by Bayer externals 7. Patients who should be information at actionable health related findings 8. External institutes request to be informed about relevant research results 9. Other restrictions (free text)	Decision body of scientists and lawyers
3. Creation of Prima4PWP group and assignment of users	PI & manager
4. Creation of empty data container in system 5. Assign usage restrictions of data in system	Appointed by decision body
6. Submit study data to container	PI alone

Conclusions

- New genomics technologies, e.g. arrays & NGS generate large amounts of data
- Analysis of genomic data has led to breakthrough treatments
- Research on patient information including their genomes is a new concept
  - Requires review of data protection regulations and new ways of working
  - *Data security and compliance is our highest priority*
- **R&D IT can deliver technical solutions to support these challenges**



Science For A Better Life

Thank you!