

Wolfgang Leister

Performance Assessment and the Human Factor

- Keynote – PESARO 2014
- Nice, February 2014

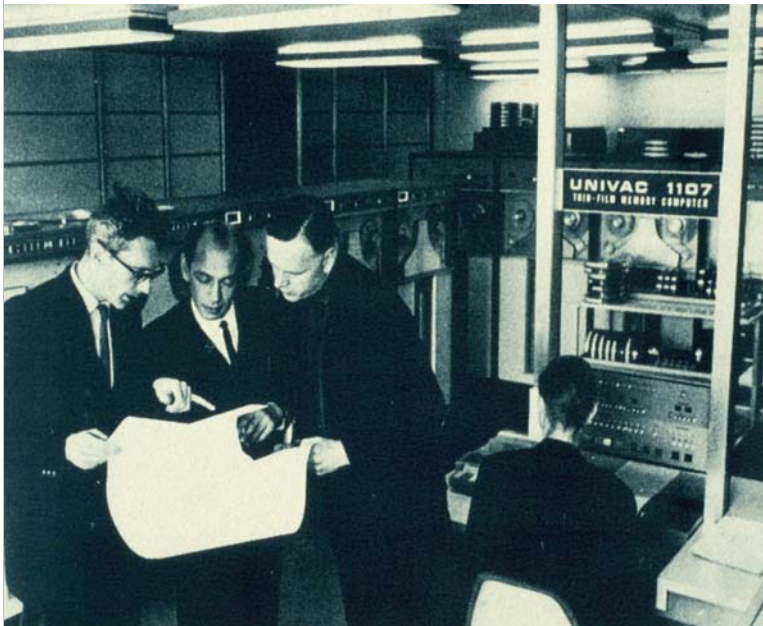


The handouts for "Performance Assessment and the Human Factor" by Dr. Wolfgang Leister are considered a collection licensed under a Creative Commons licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The parts of this collection may be licensed differently, as noted.

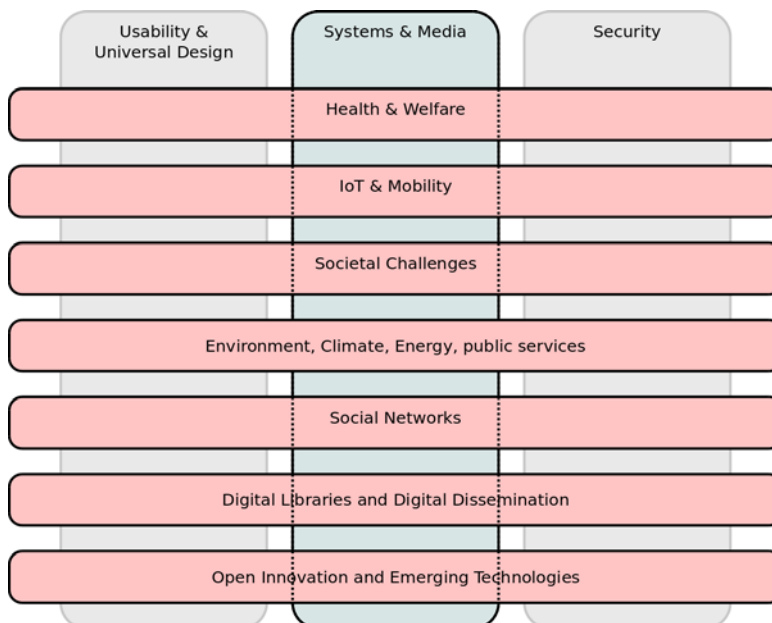


Norsk Regnesentral works with applied research. Established 1952. Ca 70 researchers. Customers are private companies, the Research Council of Norway, EU, public sector, international companies, ...

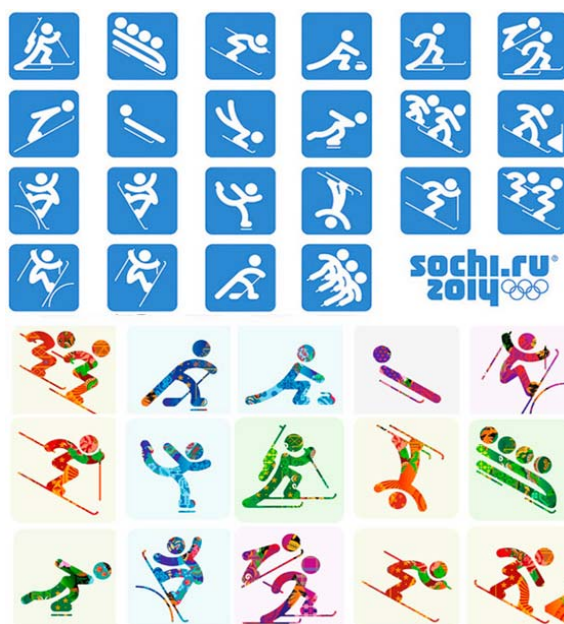
Object orientation and SIMULA



Smart Information Systems



Performance Assessment and the Human Factor



SOCHI 2014



Cross-Country Skiing

What went wrong for Norway here?

Usually they win this exercise ...

Idretter Program og resultater Medaljer Utøvere Land

Langrenn Menn 4x10 km stafett

Min tid | Arenatid

Stafett - Laura skisenter

Medaljevinnere

1	2	3
Sverige	Russland	Frankrike

Menn 4x10 km stafett Resultater

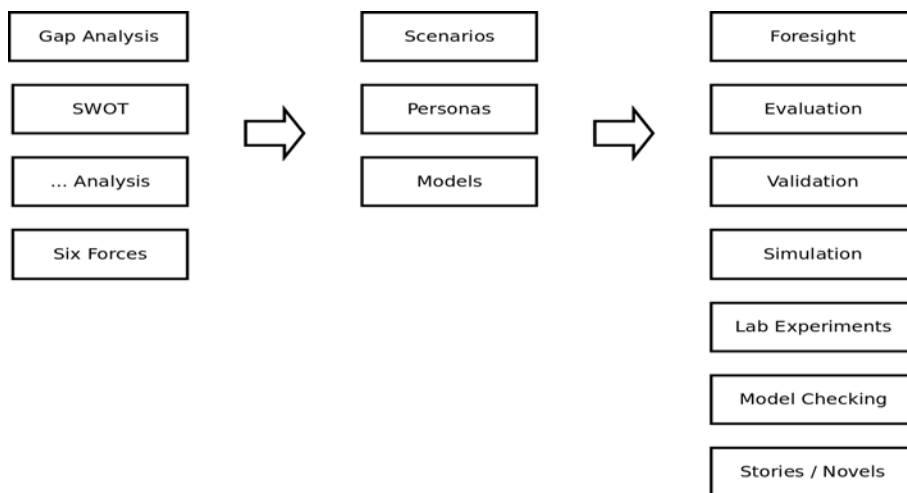
Rang.	#	Lag	Tid
1	2	Sverige	1:28:42.0
2	3	Russland	1:29:09.3
3	9	Frankrike	1:29:13.9
4	1	Norge	1:29:51.7
5	4	Italia	1:30:04.7
6	5	Finland	1:30:28.4
7	6	Sveits	1:30:33.8
8	11	Tsjekkia	1:30:36.8
9	7	Tyskland	1:31:18.8

Dommer

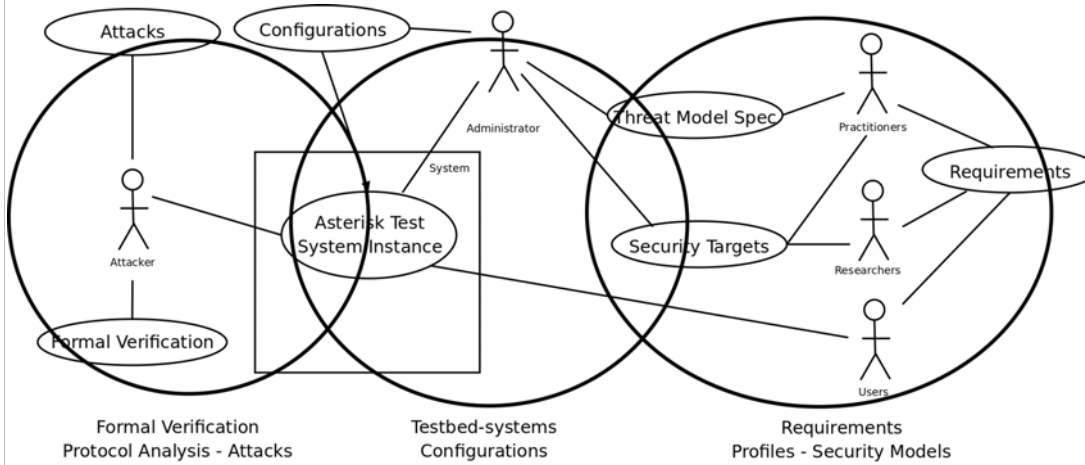
We'll come back to this later ...



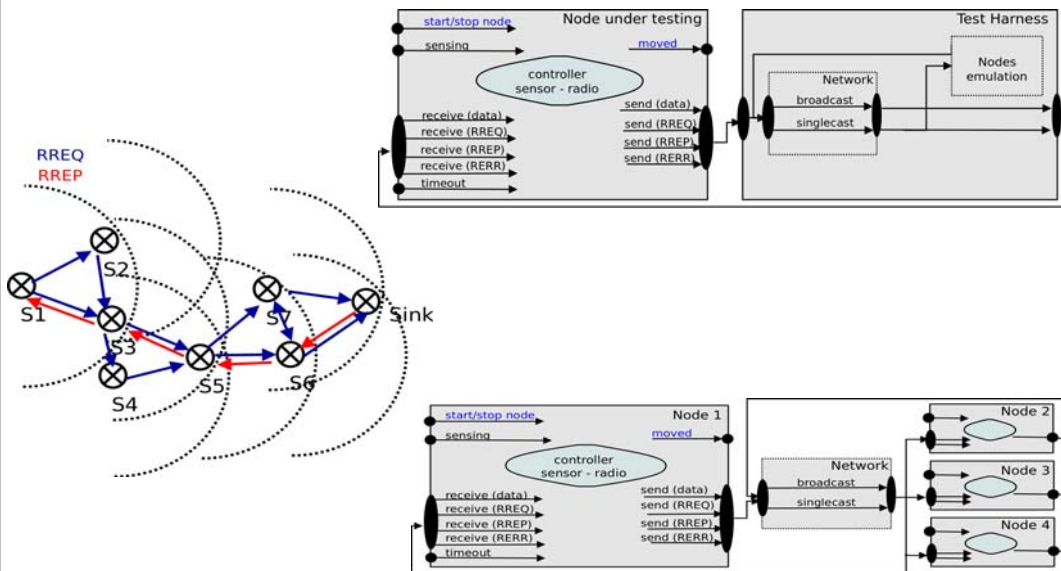
Scenarios, personas, and models, derived from thorough analyses, can be used in various prediction and analysis techniques.



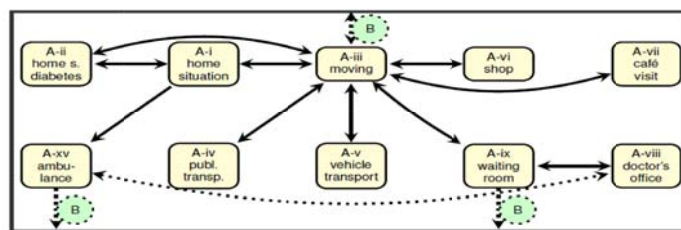
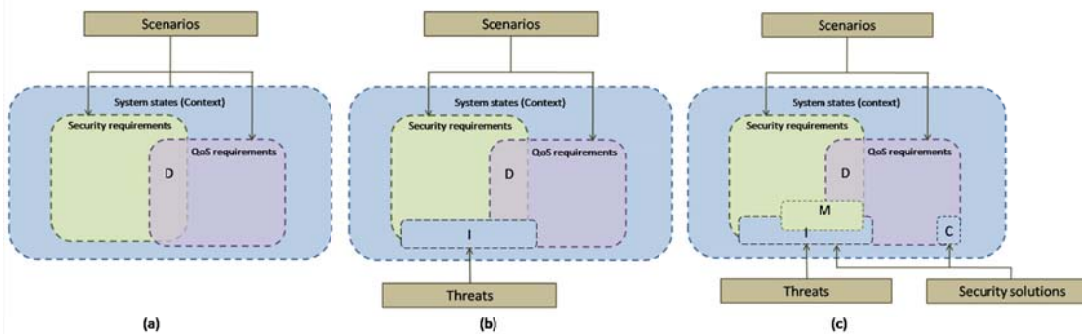
Systems can be analysed by a combination of formal analysis, testbeds, simulations, etc. - here from the EUX2010sec project



We can use models and simulations for performance assessment – here using the Credo-methodolgy



The ASSET assessment framework uses Scenarios for the evaluation.



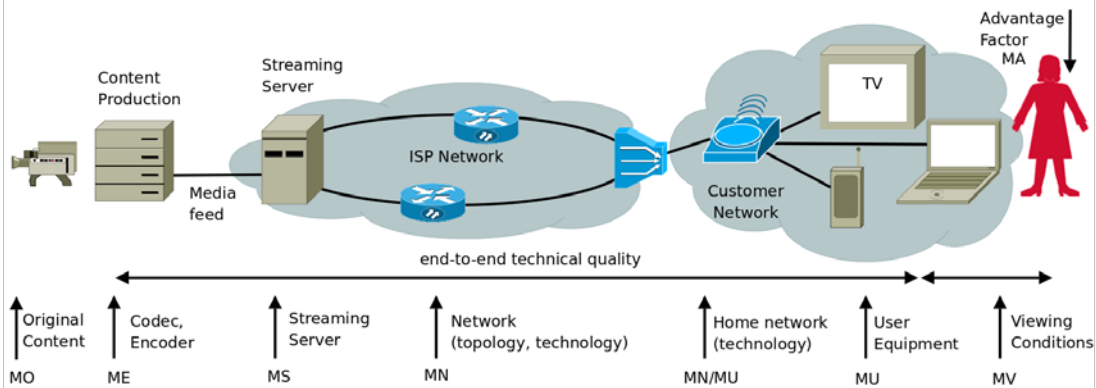
How much performance do we need ?

What about the Quality of Experience ?

User Experience !



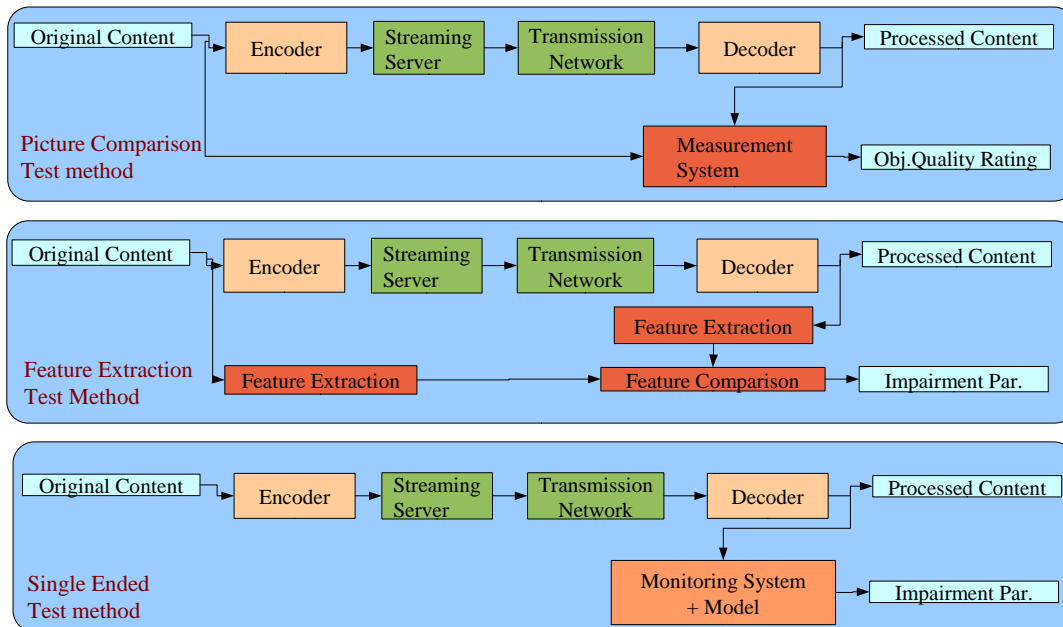
What is the performance of a video streaming system ?
And how to assess this ?



Quality estimation:

$$\tilde{Q} = Q_0 \cdot \prod_{i \in \{E, S, N, U, V, A\}} M_i$$

Classification of test methods for quality of video transmission



Assessment methods can be classified in various dimensions, such as:

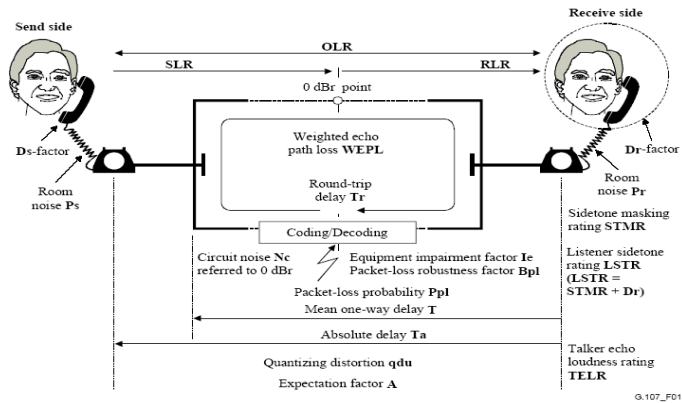
user – Subjective \longleftrightarrow vs \longrightarrow Objective – technical
picture – Directly \longleftrightarrow vs \longrightarrow Indirectly – signal
In-service \longleftrightarrow vs \longrightarrow Out-of-service
Real-time \longleftrightarrow vs \longrightarrow Deferred time
Continuous \longleftrightarrow vs \longrightarrow Samples
intrusive \longleftrightarrow vs \longrightarrow non-intrusive

QoS metrics

- carefully specified quantity related to the performance and reliability of the service that we would like to know the value of.
- Networking QoS metrics
 - Connectivity, one-way delay, two-way delay, throughput, loss, jitter, ...
 - IETF IPPM (IP Performance metrics)
RFC 2330, RFC 2680, RFC 2681, IPPM, WOAMP, ...
- Subjective Assessment Methods
 - DVB: ETR 290, TR 101 290 (Measurement guidelines for DVB systems), TR 101 291, ...
 - BT.500, BT.700 (SAMVIQ), ...

Examples of Quality Metrics: G.107 / E-Model for audio / IP-telephony

- Uses impairment factors
- Related standards:
 - G.108
 - G.113
 - G.175
 - G.562
 - ...



Examples of Quality Metrics: VQM – Video Quality Metric

- Objective measurement method
- How and what to measure
- Formulas to calculate VQM value
- Relation VQM ↔ subjective

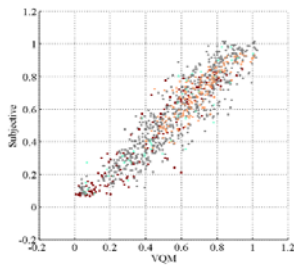
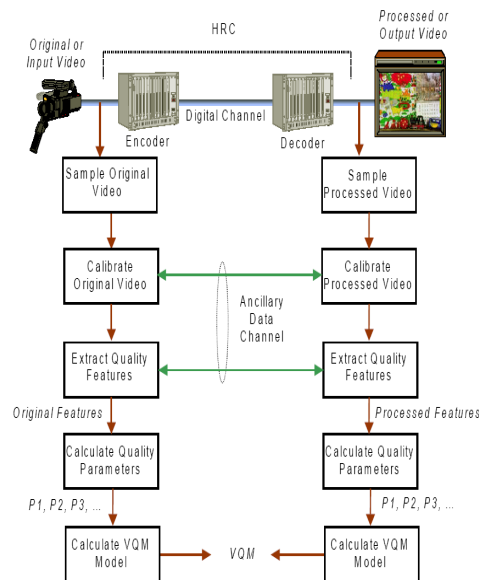


Figure 32. Clip subjective quality vs. clip VQM₀

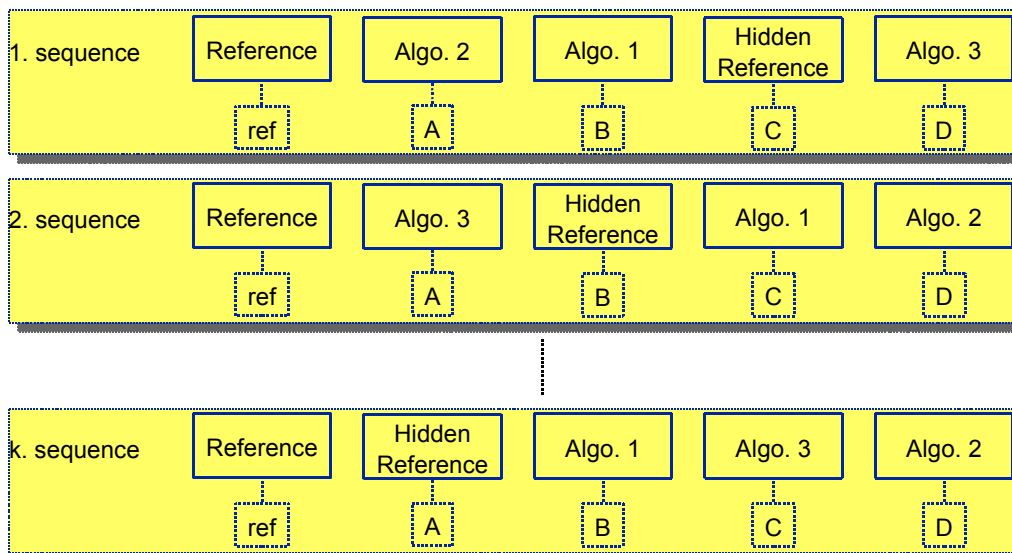


SAMVIQ / MUSHRA etc. are examples of subjective assessment methods

- Standard for Video Quality Assessment in **multimedia** (BT.700)
- Builds on experiences from **MUSHRA** (audio quality assessment)
- SAMVIQ and MUSHRA are subjective assessments using test panels.
- Uses hidden reference, low anchor, user knows total scale, possibility to change vote, scale from 0 to 100.
- High reliability, comparability w/ other test labs, objective tests, etc.

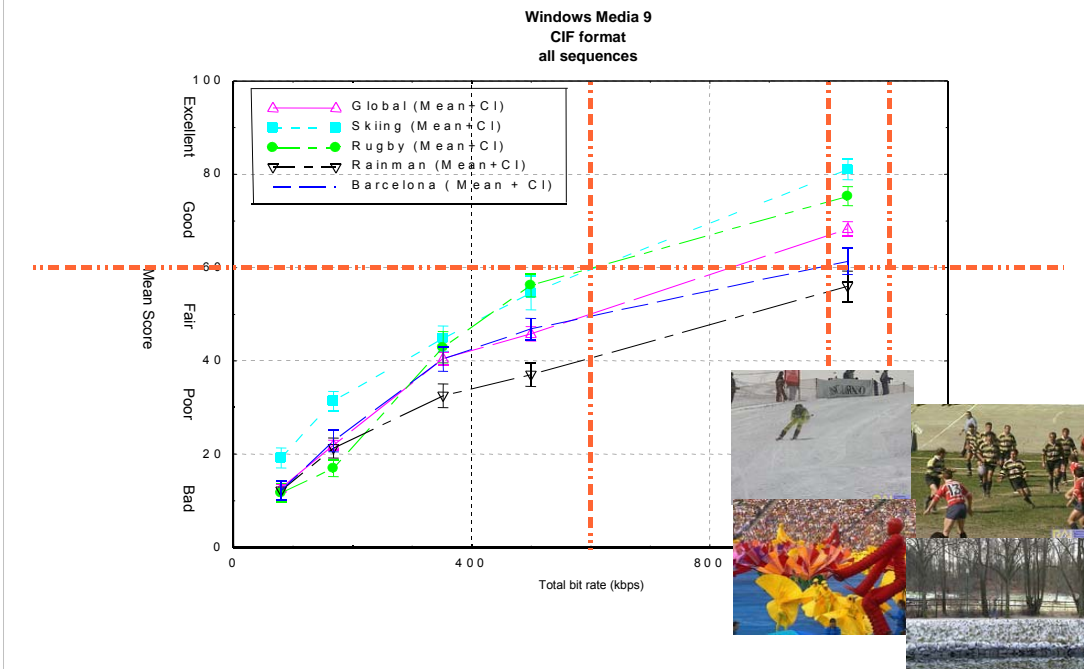


SAMVIQ: Structure of test sessions



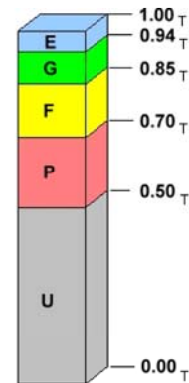
Example: Algo. 1: WM9, CIF, 168kbps
Algo. 2: WM9, CIF, 1032kbps

The content may influence the quality of a video stream

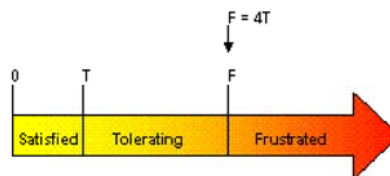


The APDEX index can be used to measure QoE for groups of subjects.

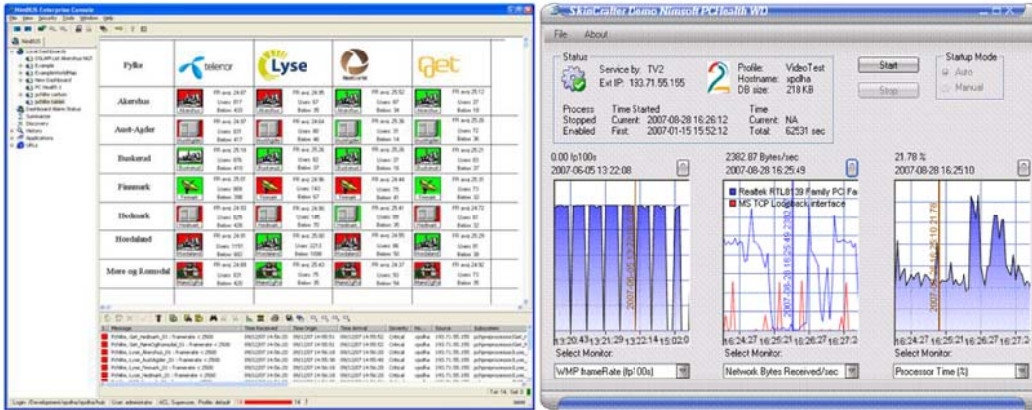
- Numerical Measure of end user satisfaction
- Ratings range from unacceptable to excellent
- Used for response times
- Three end-user categories
 - #Satisfied
 - #Tolerating
 - #Frustrated



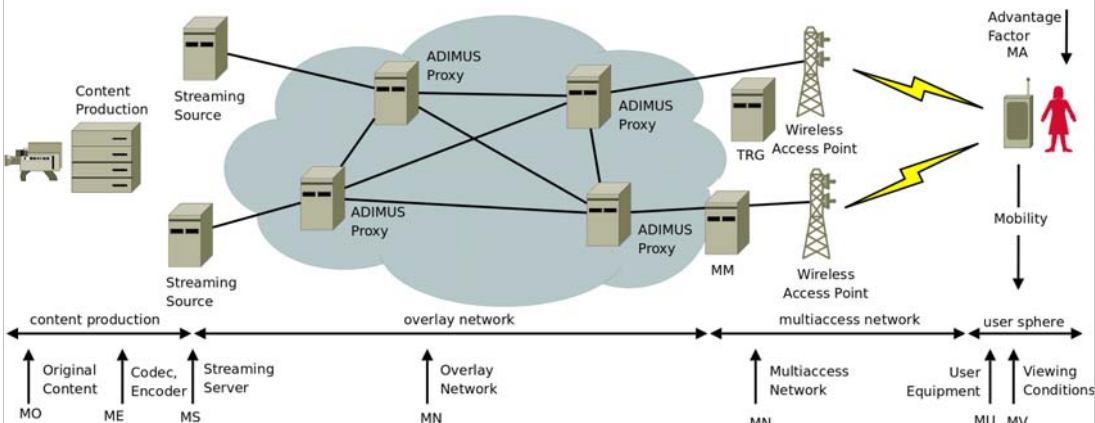
$$Apdex_T = \frac{\text{Satisfied count} + \frac{\text{Tolerating count}}{2}}{\text{Total samples}}$$



The QoE estimate of video streaming from objective measurements has been implemented in a product.



The objective and subjective assessments can be used to manage streaming systems



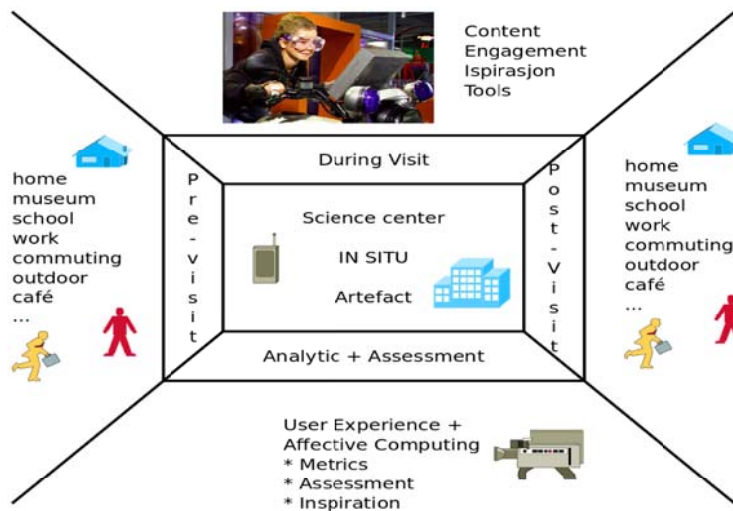
How can we assess the performance of installations in a science center ?

How can we assess engagement ?

Can we assess learning ?



The VisitorEngagement project wants to assess engagement in science centres by using sensors, cameras, data from installations, etc ...



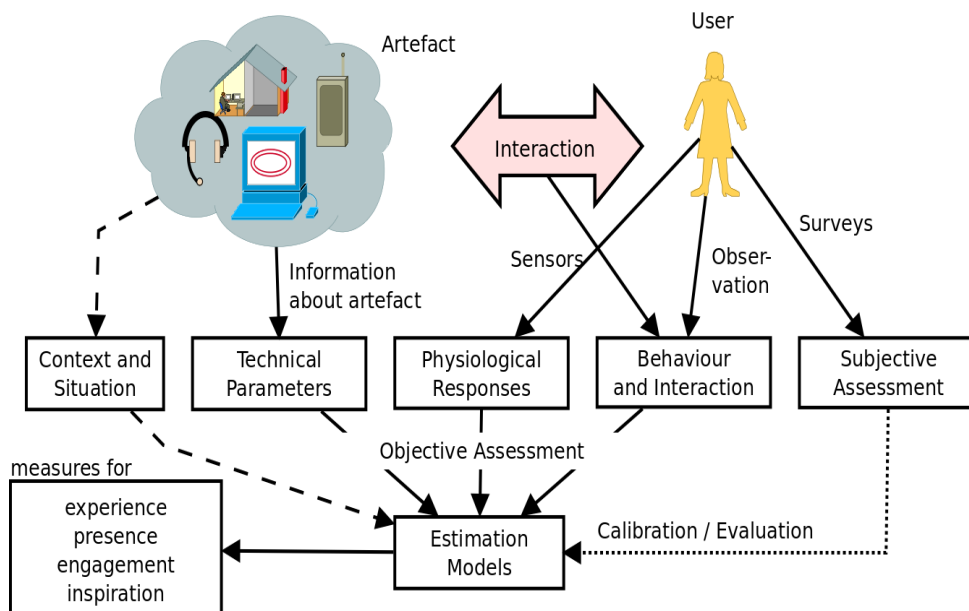
Reference to the movie «Kitchen stories» by Bent Hamer (2003):
 The project VisitorEngagement will replace Folke Nilsson by a Kinect ;-)



© SF Norge AS



We can assess objective (technical, physiological, behaviour, interaction) and subjective parameters

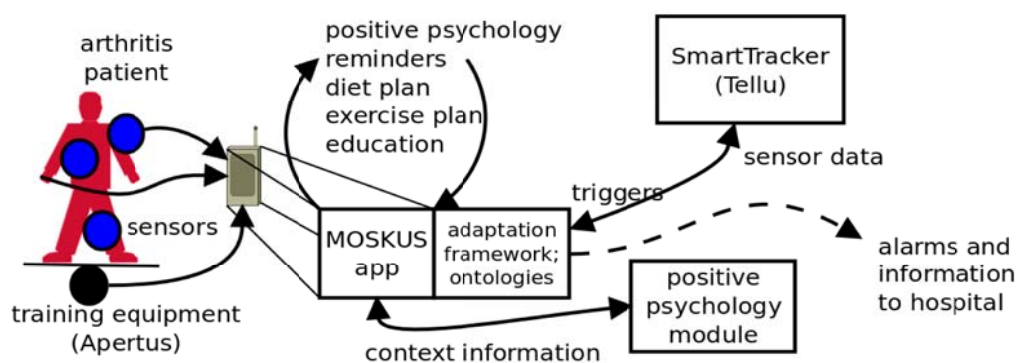


There is a variety of sensors that can be used: camera, kinect, brain signals, pulse belt, Eye-tracker, blood pressure, skin conductivity, ...



Can we apply performance measurement to health care ?

The MOSKUS project will develop smart ICT solutions to support self-management using situational feedback for arthritis, a prevalent and debilitating chronic disease.



For chronic diseases two types of assessment are used:
 (1) clinical assessment by health personnel;
 (2) subjective assessment, i.e., questionnaires, filled out by patient.

The Internet of Things
 offers new
 Opportunities !



3.2 BASFI

1. Ta på strømper eller strømpebukser uten assistanse eller ved bruk av hjelpemidler (for eksempel strømpepåkretter)	0	1	2	3	4	5	6	7	8	9	10
Lett <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Umulig											
2. Bøy deg fremover fra midjen for å plukke opp en penn fra gulvet uten å bruke et hjelpemiddel	0	1	2	3	4	5	6	7	8	9	10
Lett <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Umulig											
3. Nå opp til en høyhengende hylle uten bruk av hjelpemidler (feks gripetang)	0	1	2	3	4	5	6	7	8	9	10
Lett <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Umulig											
4. Reise deg fra en spisebordstol uten armlener eller annen hjelp	0	1	2	3	4	5	6	7	8	9	10
Lett <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Umulig											
5. Reise deg opp fra liggende stilling på gulvet uten hjelp	0	1	2	3	4	5	6	7	8	9	10
Lett <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Umulig											
6. Stå oppreist uten støtte i 10 min. uten å få ubehag	0	1	2	3	4	5	6	7	8	9	10
Lett <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Umulig											
7. Gå opp 12-15 trappetrinn uten å bruke rekkeverk eller gåstotte. En fot på hvert trinn.	0	1	2	3	4	5	6	7	8	9	10
Lett <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Umulig											
8. Se deg over skulderen uten å vri kroppen	0	1	2	3	4	5	6	7	8	9	10
Lett <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Umulig											
9. Utføre fysisk krevende aktiviteter (for eksempel fysioterapiøvelser, hagearbeid eller sport)	0	1	2	3	4	5	6	7	8	9	10
Lett <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Umulig											
10. Utføre en hel dags aktiviteter enten hjemme eller på arbeid	0	1	2	3	4	5	6	7	8	9	10
Lett <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Umulig											

Collin, A. Genet. Set et al. A new approach to defining functional ability in ambulatory spina/itis: The development of the Bath Ambulating Spinal/itis Function Index (BASFI). J Rehabil Med 2004, 11:231-5
 Oversettelse/videregående til norsk: MSD-Norge, upublisert

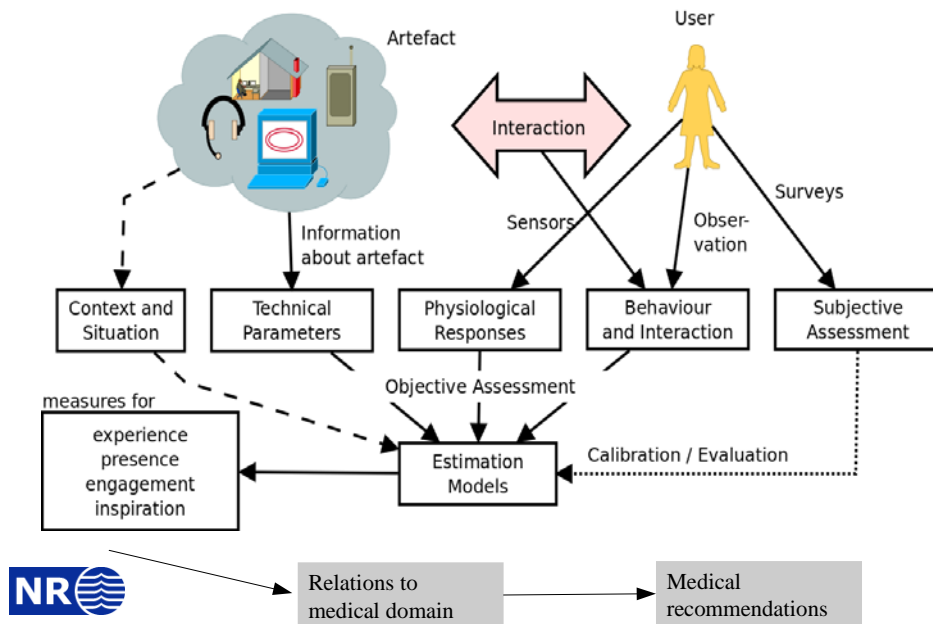
This is NOT the vision of the final result of how to assess the patient's performance.



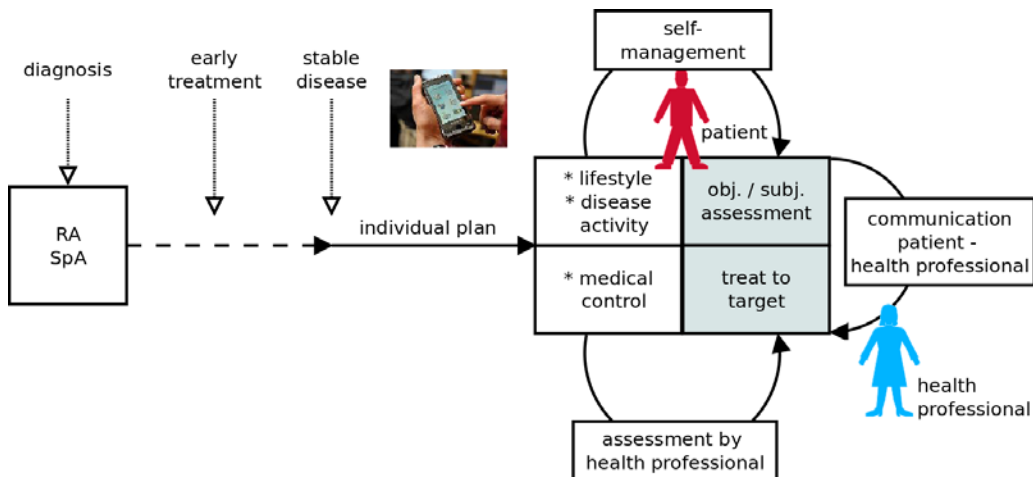
© NASA, images are in the Public Domain



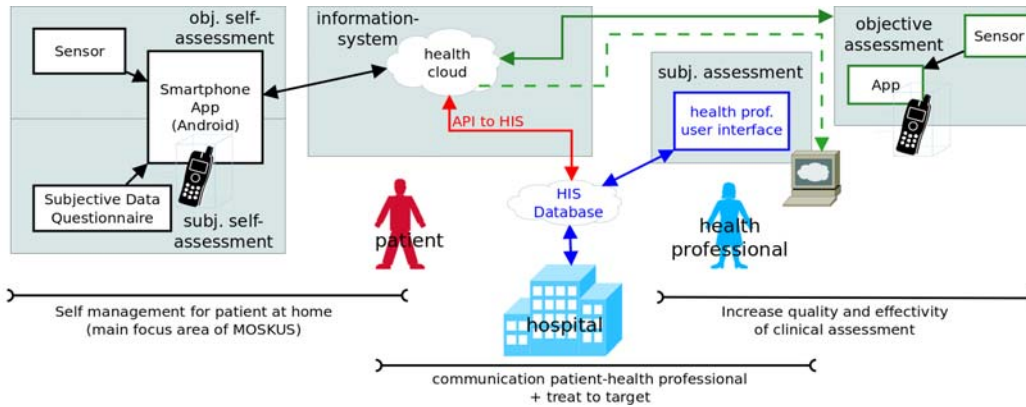
The MOSKUS project uses objective and subjective assessment techniques tailored for patients with arthritis.



The «Treat to Target» concept realised in MOSKUS: the concept supports self-management, communication patient-doctor, and assessment by doctor



Preview of the MOSKUS architecture



Assessment exercises using sensors enables 1) self-evaluation; 2) more precise communication patient-doctor; 3) tools for the doctor

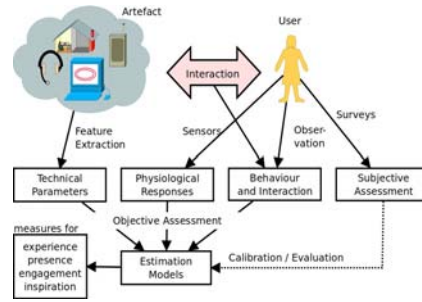


3.2 BASFI

1. Ta på strømper eller strømpebukser uten assistanse eller ved bruk av hjelpemidler (for eksempel strømpeplåtter)	0	1	2	3	4	5	6	7	8	9	10	Umulig
2. Bøye deg fremover fra midjen for å plukke opp en penn fra gulvet uten å bruke et hjelpemiddel	0	1	2	3	4	5	6	7	8	9	10	Umulig
3. Nå opp til en høyhengende hylle uten bruk av hjelpemidler (f.eks gripetang)	0	1	2	3	4	5	6	7	8	9	10	Umulig
4. Reise deg fra en spisebordstol uten armlener eller annen hjelp	0	1	2	3	4	5	6	7	8	9	10	Umulig
5. Reise deg opp fra liggende stilling på gulvet uten hjelp	0	1	2	3	4	5	6	7	8	9	10	Umulig
6. Stå oppreist uten støtte i 10 min. uten å få ubehag	0	1	2	3	4	5	6	7	8	9	10	Umulig
7. Gå opp 12-15 trappetrinn uten å bruke rekkverk eller gåstøtte. En fot på hvert trinn.	0	1	2	3	4	5	6	7	8	9	10	Umulig
8. Se deg over skulderen uten å vri kroppen	0	1	2	3	4	5	6	7	8	9	10	Umulig
9. Utføre fysisk krevende aktiviteter (for eksempel fysioterapiøvelser, hagearbeid eller sport)	0	1	2	3	4	5	6	7	8	9	10	Umulig
10. Utføre en hel dags aktiviteter enten hjemme eller på arbeid	0	1	2	3	4	5	6	7	8	9	10	Umulig

Côté, A. Gagné, S. et al. A new approach to defining functional ability in ambulating spina/limb. The development of the Bathing, Dressing, Shopping, Functional Status (BDS) scale. *International Journal of Rehabilitation Research*, 1994, 17, 229-33.





Sensors and performance assessment methods can also be applied to assess the performance of athletes.



Regarding the Cross-Country Fiasko – they are still assessing ...

Possible causes:

- Preparation of skis
- Ski wax
- The pair of skis
- ...



The New York Times

OLYMPICS Sweden's Cross-Country Skiing Success Rattles Norway

Sochi 2014 News Results Photos Interactives Follow @SochiNYT

Get daily recaps

Sweden's Cross-Country Skiing Success Rattles Norway

By KEN BELSON FEB. 18, 2014

Sweden has been beating Norway in several cross-country skiing events, including the all-important team relays. Chang W. Lee/The New York Times