

# За Научно-ориентиран AI & не слуга на бизнеса

---

**Cristiano Castelfranchi**

*Institute for Cognitive Sciences and Technologies – Roma*



# Ние (ще) живеем в **РАЗШИРЕН** и **СМЕСЕН** свят/“реалност”

Не просто "Onlife" в МРЕЖАТА (Fioridi); живеейки  
„свързано”,

**но и в нов материален свят/реалност.**

Ние ще **действаме** във Виртуалното за промяна на  
Реалното; и обратното.

Ние „**присъстваме**” където „не сме”;

виждаме и **действаме** където „не сме”.

И „някой”, който не е „тук”,

всъщност ще **действа** тук и ще „**присъства**” тук.

Ние (ще) живеем в **ХИБРИДНО**  
**общество,**

МИКС ОТ ЧОВЕШКИ И ИЗКУСТВЕН ИНТЕЛЕКТ,  
не само **Роботи**, но и **интелигентни**  
**софтуерни Агенти** или **Агенти** в нашите  
умни среди (къща, офис, кола,..)  
и *нашите когнитивни протези.*

**AI** не просто изгражда **нова технология**, а **нова Социално-Когнитивна Техническа Система**, нов свят и нова форма на общество,  
това е **антропологична революция**

**Вие сте социални инженери.** Наясно ли сте?

**Вие сте социални инженери.**

Наясно ли сте?

Ще се съсредоточа върху

**А)** Значението на **НАУЧНАТА** страна на AI;

**В)** **някои проблеми и опасности** от Дигиталната революция и от „смесената“ (виртуална и физическа) реалност и „хибридно“ общество (естествен и изкуствен интелект), в което ще живеем.

**A**

# **За Научно-ориентиран AI**

**Удоволствието от изследването (също и в AI)**  
трябва да бъде преди всичко знание, открития,  
идеи, а не само приложение и технология.

AI има твърде силна „технологична идентичност“ повече от НАУЧНА идентичност.

AI предоставя концептуални и когнитивни (формални) инструменти за моделиране и по този начин РАЗБИРАНЕ на умове, интелигентност, действие и взаимодействие, емоции, организация, ...знание.

AI трябва да се гордее с решавания принос, който донесе за **научната революция** през XX и XXI век поради **въздействието на *Науката за Изкуственото* върху поведенческите и социалните науки** (Herbert Simon)

## В науката

икономическите, социалните, техническите резултати трябва да бъдат главно „странични/непреднамерени“ ефекти. Очевидно трябва да има изследване, което не е общо ориентирано към К („основно“), а **ориентирано към решаване на проблеми**, но също така и в това „приложно“ изследване *приоритетът е К, разбиране, обяснение, моделиране..*

AI понякога изглежда малко изопачен в пълното обслужване на бизнеса, за предоставяне на нови пазарни продукти:

новото богатство, новия индустриален капитал: (Google, Amazon, и т.н.)



# Научните предимства на ИЗКУСТВЕНИЯ, **СИНТЕТИЧЕН ПОДХОД**

Към Съзнанието и Обществото

Е **РАЗБИРАНЕТО** *чрез ИЗГРАЖДАНЕ и СИМУЛИРАНЕ*

Групата ISTC CNR използва това в няколко домена. За езика, за автономията, сътрудничеството, социалността, доверието, емоциите, нормите, властта и т.н.

=====

**Научни модели на AI:**

- 1) за **моделиране, обяснение на човешкия и естествения Интелект;**

(Grosz върху Разговори & споделени планове, Ferrari's cit. Winograd)

- 2) за **подражанието** им;
- 3) **засъздаване на нова интелигентност**  
и нейната теория (“Обща Интелигентност”)

Философите често твърдят, че това, което AI и когнитивните учени правят, е да **„антропоморфизират“ машини**

(които по принцип не могат наистина да имат „ум“, „интелигентност“, „намерения“ и т.н., а просто да ги „симулират“)

За този дебат вижте например Floridi и Sanders

Точно обратното е: това, което правим, е да **„деантропоморфизираме“** такива концепции, правейки ги вече *не „антропоцентрични“, а по-обща и абстрактни*, и по-ясни, формализирани и „операционализирани“. Вече няма „думи“ на здрав разум.

Мисията на AI не е просто да *купува акритично концепции и теории от хуманитарните и социални науки и философия за тяхното „прилагане“*.

AI дава обратно решаващ принос, не само „технологичен“, като **променя тези концепции, модели и теории.**



Не само нашата околна среда и обществото ще бъдат хибридни и разширени, **НО НАШИЯТ МОЗЪК И УМ ЩЕ БЪДАТ РАЗШИРЕНИ**, нова когнитивна сила и **нови функции**.

Когнитивните ни способности не просто ще бъдат подобрени, но и **променени**.

Това не е само въпрос на „мнемонично функциониране“, външна памет, достъп до Данни и обработка, на „четене“, на „учене чрез правене“.

Ще има сериозна еволюция на нашето **„социално познание“** в Хибридното общество.

По-специално WEB („Minds on Line“) и Виртуалната реалност ще дадат възможност на:

- >> **“колективен интелект и решаване на проблеми”**,
- >> **“колективно осмисляне”**,
- >> **“капитал на знанието и споделяне”**,
- >> **“творчество”**, и

> **ново „въплъщение** на нашите когнитивни представи”

> нашето възприятие за ***пространство, време, интелигентност***,... ще бъде променено

> **изключително „външно/разпределено** познание и ум”

(P.Smart, R. Clowes, R. Heersmink: “Minds on Line”, 2017)

Една от основните функции на мозъка е **интегрирането** и **разширяването** на възприеманата реалност:

- > С възможностите, ..
- > С миналото, бъдещето (очаквания, прогнози, цели,.....)

**В**

# Революцията на **AI**: овластяване на кого?



# ИЗКУСТВЕНА СОЦИАЛНОСТ?



Co-financed by the European Union  
Connecting Europe Facility

This Master is run under the context of Action  
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom  
under GA nr. INEA/CEF/ICT/A2020/2267423



Ние - общността на AI & MAS - сме отговорни за въвеждането на „Агенти“ като

★ **“автономни”** (проактивни, с инициатива, с тяхното собствено обучение, разсъждения, еволюция, .. ) и **„социални”**,

сътрудничество с човека чрез спазване на истинските

★ **„норми”** (но също - в такъв случай - нарушаването им), и критично приемане на нашите цели (**не само**

★ **„изпълнение“**), *с прекомерна помощ, критична помощ,*

.....

И това беше правилно и неизбежно решение за истинска „Интелигентност“, взаимодействаща с нас и използвана от хората.

Ние (ISTC group) **ИЗОБЩО** не се разкайваме, че сме допринесли за модела

## **ИЗКУСТВЕНА социалност**

напротив...                      даже

Това задължава учените да **ОСЪЗНАЯТ** евентуалното **присвояване** на техните творения, на евентуални **неприемливи употреби** на тези инструменти.

## **Липсва ли ни контролът?**

**Не** за нашите автономни агенти, роботи и т.н.

а на техните **ВЪЗМОЖНИ употреби?**

Готови ли сме за **АНТРОПОЛОГИЧНАТА РЕВОЛЮЦИЯ**, основана на Intelл  
Технологиите и изкуствено смесено общество;

Кое то също е

**и икономическа, социална и политическа революция.**



Има ли **ОПАСНОСТ** от *живота* с Агенти с  
Изкуствен Интелект и Роботи?

Да бъдем **заменени** (практически или  
когнитивно) или подкрепяни и  
ръководени от тях;

Има ли **ОПАСНОСТ** в РАЗШИРЯВАНЕТО  
НА нашия ИНТЕЛЕКТ и промяна на  
КОГНИТИВНАТА ОБРАБОТКА;

## За медиите основните **ПРОБЛЕМИ** са:

- *Поверителност*
- *Сигурност (в WEB, ... при достъп...)*
- *Фалшиви новини, мисинформация*
- *Хакерски атаки*
- *Антропоморфизъм*
- *Война и изкуствени войници/оръжия*
- *Етика **вътре** в Изкуствените създания и алгоритми*

# За медиите основните **ПРОБЛЕМИ** са:

- **Война и Изкуствени войници/оръжия**

Subra Suresh, президент на Carnegie Mellon, каза, че **вкарването на етични дискусии в AI беше необходимо** с напредването на технологията. Въпреки че идеята за роботи тип „Терминатор“ все още изглежда пресилена, **армията на Съединените щати изучава автономни оръжия, които могат да вземат решения за убийство сами...**—

Най-накрая се реши проблема с бедния генерал:



## Bertolt Brecht (1898-1956)

“Генерале, вашият танк е мощно превозно средство.  
Разбива гори и смазва сто души.

Но има един дефект:

Има нужда от шофьор.

Генерале, вашият бомбардировач е мощен.  
Той лети по-бързо от буря и носи повече от слон.

Но има един дефект

Има нужда от механик.

Генерале, човекът е много полезен.  
Той може да лети и **може да убива.**

**Но той има един дефект.**

**Той може да мисли.”**

---

Най-накрая генералите вече нямат нужда от (човешки) шофьор  
или механик !! **Шофьорът с AI може да мисли, да; но ние  
генералите можем да решаваме и контролираме КАК ще  
мисли! (?)**



Co-financed by the European Union

Connecting Europe Facility

This Master is run under the context of Action  
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom  
under GA nr. INEA/CEF/ICT/A2020/2267423



# “Инженерство на Морални Агенти” :

Dagstuhl Seminar etc.

“Напълването на роботите и автономните системи с **етични норми и ценности** е все **по-неотложно предизвикателство**, като се има предвид бързото развитие, например в автомобилите без водачи, беспилотните летателни апарати (дронове) и роботите-асистенти за грижи.”

➤ *прилагане на морални разсъждения и поведение в автономни системи*

➤ **НЕ само наблюдение, но ИНТЕРНАЛИЗИРАНИ ценности и контрол**



## За медиите основните **ПРОБЛЕМИ** са:

- *Поверителност*
- *Сигурност (в WEB, ... при достъп...)*
- *Фалшиви новини, мисинформация*
- *Хакерски атаки*
- *Антропоморфизъм*
- *Война и изкуствени войници/оръжия*
- *Етика **вътре** в Изкуствените създания и алгоритми*

## За мен не по-малко сериозни проблеми.....

Оставянето настрана на бъдещето на **РАБОТА** в икономика 4.0!

**B1**

Дали нашето изследване на интелигентните  
технологии е **САМО БИЗНЕС ОРИЕНТИРАНО**  
**просто защото се нуждае от пари?**

## Среща на умовете за машинния интелект

**Лидери в индустрията, компютърни учени и студенти, и рискови капиталисти** се събират, за да обсъдят **как по-умните компютри преустройват нашия свят.**

След като една машина бъде обучена, тя може да помогне на експертите да вземат по-добри решения

... **опитни машини могат да ни помогнат да оценим (социалните) политики. И т.н....**

Само **ТОВА ЛИ СА ПРАВИЛНИТЕ СУБЕКТИ/УМОВЕ КОИТО ДА ВКЛЮЧИМ?**

за обсъждане на етични, политически и социални последици от машинния интелект и хибридно общество?

Ами **другите субекти**, които могат да бъдат включени: морални и политически философи, социални учени, профсъюзи, социални движения (като женското движение, като „окупирай Уолстрийт“...), политици, бедни страни и т.н.?



Среща на умовете за машинния интелект

Лидери в индустрията, компютърни учени и студенти, и рискови капиталисти се събират, за да обсъдят *как по-умните компютри преустройват нашия свят.*

Защо съюза е само между **академията, учените и капиталистите и бизнесмените?**;

Толкова ли е **ОЧЕВИДНО** и **БЕЗСПОРНО** в сегашната култура, че да стане **НЕВИДИМО?**

Среща на **умовете** за машинния интелект

**Лидери в индустрията, компютърни учени и студенти, и рискови капиталисти** се събират, за да обсъдят **как по-умните компютри преустройват нашия свят.**

След като една машина бъде обучена, тя може да помогне на експертите да вземат **по-добри решения**

... опитни машини могат да ни помогнат да оценим **(социалните)**  
**ПОЛИТИКИ.** И т.н....

**„По-добре“ за кого?**

**Това не е „технически“ проблем, а политически. „По-добро“ за бедни и безсилни хора/държави**

**или за доминиращи класи, лобита, сили, държави?**



**„По-добре“ за кого?**

Това не е „технически“ проблем, а политически..

**НИЕ искаме “полезен” AI, но... за кого??**

Не приемайте, че *ако нещо е от полза, то е от полза за всички.*

В обществото има сериозни контрасти на интереси и цели. Следователно, ако нещо е полезно за X (тоест благоприятства неговите/нейните цели или интереси), е вредно за Y.

Ако AI е подчинен и полезен за печалбата и бизнес интересите НЕ е непременно от полза за работниците.

Ако е от полза за доминиращите страни, не е задължително да е от полза за бедните и колонизираните страни.

>> За да бъде **ПОЛЕЗЕН**, AI трябва първо да избере на коя страна да бъде.

# AI може да бъде МНОГО полезен

-за ДЕМОКРАЦИЯТА,

-за добър пазар, с намалени измами и манипулации;

-за социално планиране и вземане на решения и политическо въображение, проекти;

-или прозрачност и контрол, участие

“AI за **СВОБОДА**” (JICAI-ECAI '18) на ХОРАТА!



# Нов Изследователски Център за **Изследване на Етиката на Изкуствения Интелект**

By JOHN MARKOFF **NYTimes** - NOV. 1, 2016

Университетът Карнеги Мелън планира да обяви в сряда, че *ще създаде изследователски център, който се фокусира върху етиката на изкуствения интелект*. Центърът по етика, наречен **K&L Gates Endowment for Ethics and Computational Technologies** се създава във време на разрастване на международната загриженост относно въздействието на AI технологии.

Това вече **доведе до набор от академични, правителствени и частни усилия** за изследване на технология, която доскоро беше до голяма степен част от научната фантастика... Peter J. Kalis, председател на адвокатската кантора, каза, че потенциалното въздействие на A.I. технологиите върху икономиката и културата направиха **от съществено значение като общество да правим обмислен, етичен избор за това как се използват софтуерът и машините.**



# “КАТО ОБЩЕСТВО”?

“набор от академични,  
правителствени и частни усилия  
”

ОТНОВО:

Защо съюз **само** между академия, учени и капиталисти и бизнесмени (и военни сили)?

Толкова ли е ОЧЕВИДНО и БЕЗСПОРНО в сегашната култура?

Нашите изследвания в Областта на Интелигентните

Технологии са **САМО БИЗНЕС ОРИЕНТИРАНИ** само защото се нуждаят от пари?



**B2**

**Скрити Интереси**

**&**

**Технология ОСЪЗНАТОСТ**



Co-financed by the European Union  
Connecting Europe Facility

This Master is run under the context of Action  
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom  
under GA nr. INEA/CEF/ICT/A2020/2267423



Сигурност, поверителност, война, етика, .. със сигурност са много уместни въпроси, върху които трябва да помислим, **НО не и най или единствените от морална и политическа гледна точка.**

>> НЕ по-малко важни са **скрити интереси, манипулиране на нас** (потребители и програмисти), експлоатация, ... **изпразване на демокрацията и т.н.**

Учените трябва да са **съзнателни**

не просто манипулирани, неосъзнати, макар и гениални слуги на тези сили и интереси.

>> **Демокрацията не е формален и дезинформиран ритуал на гласуване.**

**НИЕ** трябва да насърчим **истинска „интелигентност“** (разбиране) и **ОВЛАСТЯВАНЕ** на хората **в/на** **еволюцията на хибридните общества.**

Не само подобрена и колективна  
**ИНТЕЛИГЕНТНОСТ**, но

подобрено и **колективно ОСЪЗНАВАНЕ**

което е решаваща форма на  
„интелигентност“,

Разбиране какво правим и **ЗАЩО** го  
правим; кой ни "бута".

ПОМОЩТА при **ВЗЕМАНЕТО НА РАЦИОНАЛНИ РЕШЕНИЯ** (чрез разкриване и коригиране на нашите рационални & емоционални **ПРЕДУБЕЖДЕНИЯ**) е ОК, но....

истинският проблем не е „нашето“ решение да бъде напълно ефективно и **рационално** (не дезинформирано или пристрастно), а:

**в полза на кого?**

С **ОСЪЗНАВАНЕ** на „**интересите**“ ние обслужваме

Интелигентните Агенти и алгоритмите трябва да ни помогнат да разберем не само Целите си и как да вземем РАЦИОНАЛНО решение (не дезинформирано или предубедено), но и да разберем

**в полза на кого?**

**Нашите „финалности“/„цели“,**

**които надхвърлят нашите умствени Цели.**

# Също и **Целите на нашите Агенти и Роботи**

Изрични и прозрачни ли са поне за нас?

---

(Ро)Ботите и Агентите трябва да бъдат *разбираеми* и *надеждни*: те трябва да могат да **ни ОБЯСНЯВАТ**:

- **ЗАЩО** правят/направиха това, което правят/направиха;
- **ПРИЧИНИТЕ** и **МОТИВИТЕ** за техните действия, решения или предложения.

НЕ ни показват техния „алгоритъм“!

Това изисква **КОГНИТИВЕН МОДЕЛ** на „причини“ и „мотиви“ за вярване и за обработка на целта и решенията. (AI за **НАУКА**)

Освен това: **Целите на нашите Агенти и Роботи** *служат за* **ФУНКЦИИ**: външни, не избрани и представени ЦЕЛИ.

---

**Дали предпочитат някакъв интерес?**

*Това прозрачно ли е за нас?*

---

На кои **ЦЕННОСТИ** отговарят?

Може би не е споделено от нас, но поне е ясно! **Или неясно?**



# ТЕОРИЯ "ИНТЕРЕСИ"

Кое е по-добро за мен и моите цели но...

... Не ги разбирам и не ги преследвам  
умишлено.

Теория за **Настойническата Роля**:

X се грижи за моите "интереси", за моето добро, дори и в противоречие с мен, с настоящите ми цели; X ми помага или ме тласка или ме задължава!

При много обстоятелства агентите ще:

- решават **вместо нас** (упълномощено или не от нас),

или

- ще ни дават препоръки или просто малко тласък (известните либерални „побутвания“) *като в маркетинга,*

**Но.. В РОЛЯТА НА НАСТОЙНИК?**

Освен това:

**Кой** преценява кое е **по-добро за мен** или за нас?

Дали това наистина е „в наш интерес“ или  
предимно в ИНТЕРЕС на доминиращи  
финансови и информационни сили?

Или (в много страни) на политическия режим?

Това важи и за по-ясно въздействащи  
устройства като

>> **ПРЕПОРЪЧИТЕЛНИ СИСТЕМИ**

КОИТО ще ни познават по-добре от нас.

Дали те ще ни дават препоръки и предложения „в наш  
ИНТЕРЕС“, в НАСТОЙНИЧЕСКО отношение,

или ще следват пазарни критерии

просто **по-ефективна, персонализирана реклама?**

От страната на „потребителя“?!

Или на „продавача“ (на нашите данни или на някаква  
стока);

Те ще **решават „вместо нас“**,

но... **ДВУСМИСЛЕНО: „вместо“** нас или също **„за наше добро“?**

---

Социалните Роботи и Интелигентните Агенти **НЯМА**  
**да управляват в свой собствен интерес**  
(научна фантастика!), а... **в чий интерес?**  
**ОВЛАСТЯВАНЕ на кого?**

И ще можем ли да наблюдаваме и  
разбираме това?

И това да **стане „прозрачно“** за  
**хората?**

Освен това: „Настойничество“ не означава да ме защитаваш”

а просто грижа само за нашите „индивидуални” “лични” интереси, но също така и помощ да разберем и да се погрижим за:

-на **Общи интереси** и възможни колективни субекти и общности и натиск;

-на скрити **конфликти на интереси**,

-на „**общото**“, на обществените блага и тяхното значение и зачитане на околната среда, енергията, водата, **общественото здраве**, ...)

# “РАЗШИРЕН ИНТЕЛЕКТ”

също означава

РАЗШИРЕНА СОЦИАЛНА ОСЪЗНАТОСТ:

КАК работи „НЕВИДИМАТА РЪКА“ (богът на либерализма)

който организира възникващия и „спонтанен“ социален „ред“.

# ПРИСЪСТВИЯ, агенти, работи, ...

Отново става въпрос:

**За кои политически и морални  
ценности ще се грижат;**

не нашият „шофьор на кола“

но „двигателите на обществото“ !!

**и нашият житейски навигатор.**



3

# Алгоритъмът "Устата на истината"



Ясно е, че разработваме **алгоритми за установяване на „истината“** в тази бъркотия от данни, от твърдения, измами и новини, които ще бъдат разпространени и достъпни чрез МРЕЖАТА.

Алгоритъм за вземане на решение за **надеждни източници, достоверна информация**, какво е „вярно“ сред толкова много различни твърдения и данни.

**Няма алтернатива за това.  
Въпреки това:**

- На каква база такъв алгоритъм ще „установи кое е вярно“? Само въз основа на надеждни и конвергентни източници? От техния брой и нет? За пряк или косвен достъп до „факта“?
- Също така въз основа на „ценностите“ и на споделянето и приемането на ценностите на източника?

Дори за „официалната“ наука: тя винаги ли улавя или казва истината?

- И ще има **догматични истини** и **безспорни авторитети**, както във всяка култура?



*The people will believe what  
the media tells them they believe.*  
- George Orwell



Co-financed by the European Union  
Connecting Europe Facility

This Master is run under the context of Action  
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom  
under GA nr. INEA/CEF/ICT/A2020/2267423



- И **коя култура и ценности** ще се приемат за „правилните“?

Как ще ни бъде позволено да разграничим конфликт на ценности или интереси от обикновен конфликт между повече или по-малко достоверни данни, повече или по-малко обосновани, директни, контролирани, надеждни, ...?



3

**‘ПРИСЪСТВИЯ’**

**в нашата**

**СМЕСЕНА РЕАЛНОСТ и ОБЩЕСТВО**

Автономните и проактивни интелигентни субекти ще станат „присъствия“ и „роли“ в нашето *хибридно общество* (човек и изкуствен агент).

и **смесена и разширена** реалност (комбинирана виртуална и „истински“, „естествен“ и автоматичен/протезен свят).

Сега проблемът ще бъде:

**можем ли да управляваме тези автономни и твърде информирани и интелигентни агенти?**

Става дума за:

➤ **A)** Какви **роли** ще играят тези материални или нематериални, видими и невидими „същности“ в нашия живот и околна среда;

(work with Ricci & Tummolini)



А) Кои **роли**

Ще бъдат ли

**Нашите**

**пазители**

с „настойническа“ роля?

Като ни помагат, защитават и ни  
дават сили

**Ангели**



## A) Кои **роли**



Или – по-малко религиозно – нашият Jiminy Cricket  
(Говорещият Щурец)  
с неговите препоръки?

Например, ние няма да бъдем адресат и мениджър на нашето  
собствено „количествено Аз“ и lifeLog“.

# A) Кои роли

Или нашия

# ръководител ICT- Panopticon

в който живеем?

The Prisoner and the Free



can't see the watcher

knows he's being watched

rounded back

all alone  
within a crowd of people



Connecting Europe Facility

This Master is run under the context of Action  
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom  
under GA nr. INEA/CEF/ICT/A2020/2267423



# А) Кои **роли**

или **нашия изкушаващ Дух**



## А) Кои роли

или нашия изкусителен

**Дявол:**

в полза на някаква  
маркетингова политика или  
монопол, или влиещ и  
манипулиращ мениджър за  
скрити политически или  
икономически правомощия?



# СМЕСЕНА РЕАЛНОСТ, СМЕСЕНО ТЯЛО И УМ

Ще ги "вградим", почувстваме ли като **части** от "нас", нашата "**умствена протеза**"?

Ще се вслушаме ли в този морален или рационален „глас“, като **наш собствен** ум или съзнание изразява **нашето (разширено)** *СуперЕго.*

# СМЕСЕНА РЕАЛНОСТ, СМЕСЕНО ТЯЛО И УМ

Или нашето **Супер Его** ще бъде  
**„екстернализирано“?**

Не „аз“.

Ще се вслушаме ли в „нея“ като в гласа на  
нашата майка, на нашата учителка?

Или ще станем „слушатели на глас“??

# СМЕСЕНА РЕАЛНОСТ, СМЕСЕНО ТЯЛО И УМ

И двете решения вероятно ще бъдат там:

>"Социалното": **Външни гласове и Агенти**. Нашият най-добър приятел; нашите сексуални партньори,...

&

>„Рефлексивно социалният“: **разширен интернализиран Аз и Съзнание**



# ПРИСЪСТВИЯ

Отново става въпрос за:

➤ **В) За кои политически и морални ценности ще се грижат?**

не нашият „шофьор на кола“, а „шофьорите на обществото“ !! и **нашият житейски навигатор.**

Те ще **решават „вместо нас“**,

но... **ДВУСМИСЛЕНО: „вместо“** нас или също **„за наше добро“?**

---

Социалните Роботи и Интелигентните Агенти **НЯМА**  
**да управляват в свой собствен интерес**  
(научна фантастика!), а... **в чий интерес?**  
**ОВЛАСТЯВАНЕ на кого?**

И ще можем ли да наблюдаваме и  
разбираме това?

И това да **стане „прозрачно“** за  
**хората?**

**B3**

# ТЕХНОЛОГИИ ЗА **НЕ**СЪГЛАСИЕ

**A) Има твърде силна идеология и реторика за **обществото като сътрудничество, общо намерение, колективни предимства,.....****

как да се постигнат удобни споразумения, равновесие и т.н.

Освен това мрежата (неслучайно) благоприятства едно отклоняващо се политическо чувство: „ние“ срещу „тях“ (управители, политическа каста, централизирани власти).

Това възприемане на „ние“ е напълно подвеждащо: няма „ние“ с общи ценности, цели и интереси, които трябва да бъдат обединени срещу политическата власт като такава (в случай срещу реалната власт (финансова власт), която е узурпирала политическата власт).

Населението се състои от **различни класи, полове, поколения, .. и култури** с много различни и противоречиви ценности и интереси;

това е **ИСТИНСКИЯТ КОНФЛИКТ**

(не „ние“ и „те“),

и политическата дейност и сили трябваше да представляват и защитават именно тези различни социални интереси, а не само „общия“ интерес.

Някои конфликти на интереси или ценности могат да бъдат решени и съгласувани в **общ интерес**, но голяма част от политическите/правителствените решения не са за **общо предимство** (с изключение на намаляване на гражданската война), за справедливо разпределение;

то е за преобладаване или насърчаване на интересите на дадена група (класа, лоби, пол, възгледи, ...) чрез намаляване на правомощията на другите.

# Необходимостта от конфликти

## Конфликти: предпоставката за Демокрация

Конфликтите **не** са просто **конфликти на възгледи или мнения**, или поради различни концепции, информация, разсъждения.

## Има конфликти на „обективни интереси“

проблемът е в конфликти между интереси на група или класи, или конфликти между "частни" интереси срещу общи интереси, "общи" и обществени блага.

Социалните конфликти всъщност нямат "вербално/когнитивно" или "техническо" решение, само базирано на **данни** и технически принципи; **те имат "политическо" решение;** това е въпрос на "власт" и на преобладаващи интереси и компромиси (равновесие, разделение/споделяне).



# Необходимостта от конфликти

## Конфликти: предпоставката за Демокрация

### *Няма конфликти, няма демокрация*

Демокрацията не е само "отговор" на Cs и за тяхното модерирание; това би било начин за насърчаване, растеж (и решаване).

Конфликтите не само трябва да бъдат управлявани, намалявани, помирявани: те дори трябва да бъдат насърчавани и това всъщност е ролята/функцията на специфични сили и организации, като синдикати, партии, групи от интереси, асоциации, движения и т.н. Решаващи заинтересовани страни на демокрацията, но и определено отговорни за типичния социален, културен, икономически "напредък" на западните страни през последните векове, а сега и на останалия свят.

Разбира се, конфликтите могат да бъдат опасни, като ни водят до битки, насилие, войни, .. Така че е вярно, че обществата и групите се нуждаят от "правила", за да ги управляват, за да избегнат дегенерация. Централизираната държава беше едно от тези решения: държавата монополизира насилието; частното или груповото насилие е забранено.

# Да живеят конфликтите!

Конфликтите

с техните **несъгласия и споразумения**

следователно са двигателят и принципът на демокрацията и на нейната възможна ефективност при промяна на обществото в полза на подчинените субекти, класове и групи в неравностойно положение и т.н.

**Да живеят конфликтите!**

# Демокрация

Марк Твен е брилянтно прав

''Ако гласуването имаше някакво значение, те нямаше да ни позволят да го направим''.

Но... проблемът е много по-труден; това не е просто заговор, а че гласуваме по саморазрушителен начин и като цяло нашата колективна глупост.

**Може ли политическото „образование“ и образованието за „обща“ § Цифровото общество и демократията на участието да са достатъчни и да решат този „когнитивен“ и социален проблем;?**

Те ще помогнат. Но

като се има предвид непосредственото местно възприемане на противоречивите интереси и конкуренцията и слепотата за общите интереси между различни страни и бедни класи и етнически групи, както и чувства за принадлежност и идентичност, конформизъм и групова психика в група срещу извън групата, ... имам някои съмнения..

След няколко века те ще видят.

- Да се **даде глас на хората**, които никога не са в състояние да протестират, и да бъдат чути...
- Да накарат **конфликтите** да се **появят** и осъзнаят, да изразят несъгласието си, да направят прозрачни кои интереси са скрити и преобладаващи, ...

трябва да бъде (в демокрацията) **една от основните задачи на интелигентните социални технологии.**

## **V) “критично мислене”**

Използването на WEB технологии за организиране на „движения” е ОК; но **не е толкова добре без насърчаване на критично съзнание**

Не само от:

- Противодействие на нашето ***Пристрастие към Потвърждението;***
- Противодействие на ***склонността ни към общителност и „ефекта на балона“*** в МРЕЖАТА

но като ни помага да

**разберем скритите сили, а също и нашите предразсъдъци.**

Нуждаем се от среда и Агенти за учене и развиване на отношение на **„критично мислене“**; да управляваме нашите **когнитивни и мотивационни пристрастия**; и т.н.

Да ни подкрепят в аргументацията и дискусията, и в разбирането на подвеждащи аргументи на другите.

Да се противопоставим на преобладаването на „аудитория“ срещу „качество“, на самомаркетинг и индекси срещу оригиналност и качество; и т.н.

**....за пропагандата, Академията, джендър моделите, фанатизма, суеверието, градските легенди,...**

**Имаме впечатляващи възможности с нови интелигентни и взаимодействащи технологии, големи данни и т.н..**

**Те не трябва да се използват само за продажба и за доминиране.**

# Демистифициране на идеологията на **NET**

*NET взаимодействието се възприема като нейерархично, без надстройка и посредничество, индивидуално управлявано, спонтанно, следователно „свободно“. Наистина и директно „демократично“.*

Неолиберален възглед и погрешно възприятие.

- Има **нови Сили** отвъд WEB и нейната дейност и информация;
- Впечатляващи **олигополни икономически интереси**
- **Влияние, манипулация,**
- **Експлоатация** на данни, експлоатация на работа

# ТЕХНОЛОГИИ ЗА НЕСЪГЛАСИЕ

## γ) **анти-манипулация**

антиманипулационни ИСТ и когнитивни технологии се използват за разпознаване на нашия профил и интереси, **НО, НЕ за ДАВАНЕ НА СИЛА**, ами

за да ни предложи/**накара да „купим“ нещо** (стоки, идеи, ...) Те ни наблюдават и анализират, за да ни манипулират и влияят на нашите избори.

---

Нуждаем се от AI технологии **против манипулация**:

Бих искал да имам не толкова личен виртуален или роботизиран психотерапевт или физиотерапевт;



# ТЕХНОЛОГИИ ЗА НЕСЪГЛАСИЕ

## γ) *анти-манипулация*

Бих искал много повече „**навигатор на живота**“ в основната си „социална роля“ (напр. потребител!), но не и навигатор, който казва „завийте надясно, завийте наляво“, „купете това; не купувайте това“ ...

Но **учител, обучител**, който ме кара да разбера и да разсъждавам защо съм ориентиран в тази посока, аз избирам този продукт; тревожа се дали имам правилната информация или имам погрешни вярвания и т.н...

Осъзнаване **кой** и **как** ме убеждава или просто **несъзнателно** ме манипулира; и така нататък..

# Заклучителни белешки



Co-financed by the European Union  
Connecting Europe Facility

This Master is run under the context of Action  
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom  
under GA nr. INEA/CEF/ICT/A2020/2267423



Голямата **РЕВОЛЮЦИЯ** на ICT, на цифровото наблюдение и прогнозиране (чрез симулация) и **ГОЛЕМИТЕ ДАННИ** може да даде на обществото (на демонстрации)

**стъкло** през което да се наблюдават и да следят какво се случва.

**Стъкло, отразяващо и това, което е невидимо: скрити присъствия, бъдеще (предсказания за планиране):**

**СТЪКЛО на НЕВИДИМОТО**

Голямата **РЕВОЛЮЦИЯ** на ICT, на цифровото наблюдение и прогнозиране (чрез симулация) и **ГОЛЕМИТЕ ДАННИ** може да даде на обществото (на демонстрации)

стъкло през което да се наблюдават в **бъдещето:**

“Най-добрият начин да изградите бъдещето е да го предвидите/симулирате/представите”

(Nala Yak)

# СТЪКЛОТО на НЕВИДИМОТО

Не само „ПРИСЪСТВИЯ“; какво „не присъства“ тук, но може да присъства виртуално за взаимодействие, може да действа в тази дума и обратно и Т.Н.

Но също и стъкло, способно да покаже това, което не може да се види/разбере: бъдещето, прогнозите, „възникващият“ ред и

**скрити явления и интереси**

(например, мога ли да **видя** кой сега получава личните ми данни? И за какво?;

За кого работя безплатно?)

# Можем ли да Преодолеем Човешкото Отчуждение?

Можем ли, като експлоатираме

>колективен, разпределен, хибриден ИНТЕЛЕКТ

и

>ГОЛЕМИ ДАННИ

и

>обратна връзка по време на работа и информация от местни заинтересовани страни и интелигентни сензори

и

>Компютърно ОБУЧЕНИЕ и ПРОГНОЗИРАНЕ

и

>Компютър (базиран на агент) Социална СИМУЛАЦИЯ и ВИРТУАЛНА РЕАЛНОСТ и СЕРИОЗНИ ИГРИ и др.

МОЖЕМ ЛИ

**ДА НАПРАВИМ ВИДИМА НЕВИДИМАТА  
РЪКА?**

**и (частично) ДА Я УПРАВЛЯВАМЕ?**

# Можем ли да Преодолеем Нашето Отчуждение?

Ще стане ли **Левиатан**

гигантска *свързана и информирана*  
*общност от агенти,*

управлявайки тяхната Колективна Сила?

1. Скептичен съм към това (също поради когнитивни  
причини)

2. Притеснявам се за възможна **нет-Демагогия**

***За да Видите Какво (в момента) е  
Невидимо:***

***Изкуствено Увеличена Осъзнатост***

**истинската революция на AI**

**(Включително себе си! То  
използва)**



В цифровото общество

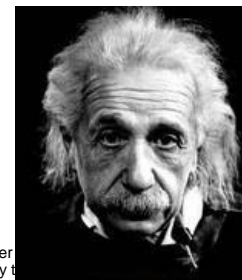
**Изкуствения Интелект**

може или да експлоатира

или да преодолява

нашата **Естествена Глупост**

“Две неща са безкрайни: Вселената и човешката глупост; и не съм сигурен за вселената.”



ИЗВИНТЕ за такъв ПЕСИМИСТИЧЕН РАЗГОВОР  
не е много смешно

НО

но ми се иска да разберете посланието на

**Оптимизмът ВОЛЯ (Gramsci)**

**и удоволствието/красотата на AI**

КРАЙ

Благодаря ви за вниманието!

Искам да **благодаря** на нашата изследователска група по  
когнитивна наука в ISTC:

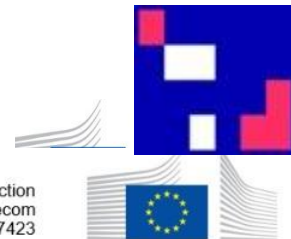
**‘групата GOAL’**

*Rino Falcone (Emiliano Lorini) Maria Miceli Fabio Paglieri Giovanni  
Pezzulo Luca Tummolini*

.....

И В ПАМЕТ на **Rosaria Conte** (*Social Simulation LABSS Group*)

<http://www.istc.cnr.it/group/goal>



Co-financed by the European Union  
Connecting Europe Facility

This Master is run under the context of Action  
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom  
under GA nr. INEA/CEF/ICT/A2020/2267423