

TNO innovation
for life

DATA FOR GOOD? EXPERIMENT!
DR. ANNE FLEUR VAN VEENSTRA

Overheid als flappentap voor criminele Rotterdammers

NADIA BERKELDER

ROTTERDAM | Criminele Rotterdammers sluisen geld weg bij de overheid, zorgverzekeraars en corporaties. Bij het eerste grote onderzoek met 'big data' naar fraude in Nederland zijn aanwijzingen gevonden dat dit op grotere schaal voorkomt.

Het onderzoek vond plaats met geanonimiseerde gegevens van mensen in Rotterdam Zuid en Delfshaven. Het was een experiment, vertelt Annemiek Roobeek, hoogleraar op Nyenrode, die het onderzoek leidde. Gegevens van mensen die

met justitie in aanraking zijn geweest, zijn naast bestanden van de gemeente, zorgverzekeraars, scholen en banken gelegd en met elkaar vergeleken. Alle gegevens zijn inmiddels vernietigd.

Patronen

Het was geen strafrechtelijk onderzoek, maar een onderzoek naar criminele patronen. „We hebben onzichtbare onveiligheid blootgelegd,” zegt Roobeek, die samenwerkte met het Openbaar Ministerie. Het onderzoek geeft de overheid en bedrijven aanwijzingen over waar ze moe-

ten zoeken naar fraude en criminaliteit. Zo zagen onderzoekers dat dertig verzekerden van hetzelfde rekeningnummer gebruikmaakten voor zorgdeclaraties, dat ergens anders tientallen mensen op één adres woonden en dat één man contant aan de balie betaalde voor tien sociale huurwoningen. „Als je dat soort dingen weet, kun je jezelf daar beter tegen wapenen,” vindt Roobeek. „We krijgen hiermee niet de boeven in beeld, maar we begrijpen wel beter waar de zwaktes in het systeem zitten en hoe het werkt.” » P3

Drugsroute in beeld door zorgdeclaraties

ROTTERDAM | De uitkomsten van het onderzoek naar fraude met 'big data' worden gebruikt om criminaliteit op te sporen, laat de gemeente Rotterdam in een reactie weten.

Dat criminelen op grote schaal geld weghalen bij de overheid en bedrijven was voor veel organisaties een verrassing, zeg onderzoekster Annemiek Roobeek. „Het gaat bij fraude misschien om 100 euro per keer, of een paar 100 euro, maar tel dat maar eens bij elkaar op. Dat is gigantisch.”

Bedrijven en organisaties stonden niet te springen om mee te doen aan het experiment en hun gegevens beschikbaar te stellen. Een van de bedrijven die wel vanaf het begin enthousiast was, is zorgverzekeraar DSW uit Schiedam. „Wij kijken intern ook naar patronen om fraude op te sporen,” vertelt bestuursvoorzitter Chris Oomen. „We willen daar beter in worden en dat was voor ons de belangrijkste reden om mee te doen.”

„We zagen een groep mensen die zorg declareerde langs een bepaalde

route. Dat bleek een drugsroute van Rotterdam naar Limburg te zijn. Voor justitie was dat een bevestiging dat zij op die route actief waren. En zo hebben we meer zaken gezien. We doen daar verder niets meer mee, we hebben alle gegevens vernietigd. We hebben alleen meegeedaan om te kijken of het iets oplevert.”

De gegevens van bedrijven en organisaties zijn, in tegenstelling tot wat mensen vaak denken, niet zo makkelijk op elkaar te leggen en te vergelijken. „Publieke organisaties hebben vaak hun dossiers niet op orde,” zegt Roobeek. „Mensen zeggen dikwijls: ze weten alles van je, maar er moet nog heel wat water door de Maas voordat het zover is.”

- › Combination of data from different organizations(e.g. municipal, health care insurance, schools, banks)
- › For privacy reasons, data was deleted afterwards
- › No criminal charges were pursued, it was treated as an experiment
- › Many instances of 'small fraud', make up a large amount
- › Administrations of public and private organizations' were often not of sufficient quality

DATA-DRIVEN POLICY

- › Data e.g. (real-time) **sensor data** may provide new insights for policy making
- › This may also require new methods for data analysis, such as **machine learning**
- › Challenges include **organizational readiness** and **policy makers' willingness** for using data and data-driven methods for policy making
- › To create support, it is necessary to involve **citizens and other stakeholders** in different phases of the policy cycle



› WHY EXPERIMENT?

- › Explore the impact of **new data sources** and **analytics** for public policy and public services
- › Support **policy makers** to make better use of such approaches and develop **tested methodologies** for data-driven policy making
- › Legal restrictions, such as the GDPR, may require a **safe environment** ('regulatory sandbox')
- › Involve citizens (and other stakeholders) in policy making early on ('co-creation')
- › Investigate the opportunities and challenges that arise for **scaling**



› POLICY LAB APPROACH

1. **Identify new data sources** and technologies that impact public policy.
2. **Design experiments** to test new technologies, methodologies and policy models.
3. **Implement and monitor** policy; develop opportunities for **scaling**.



- Methods & Models
- Experimentation & Co-creation
- Impact & Learning

YOUTH POLICY MUNICIPALITY OF ROTTERDAM

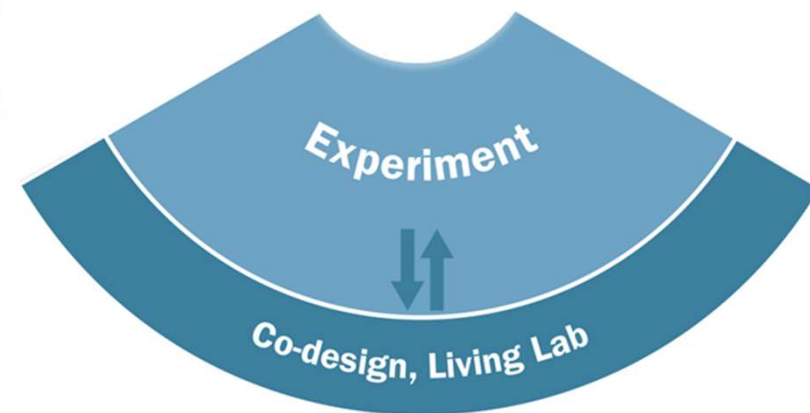
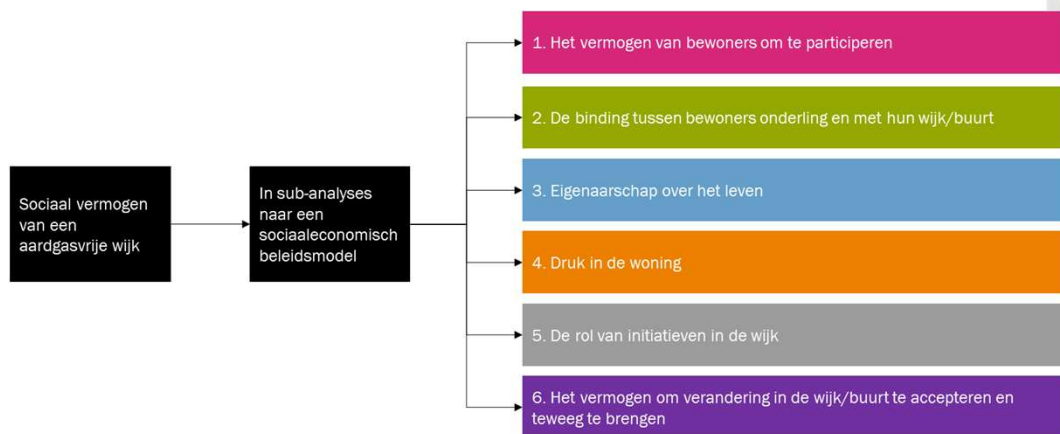
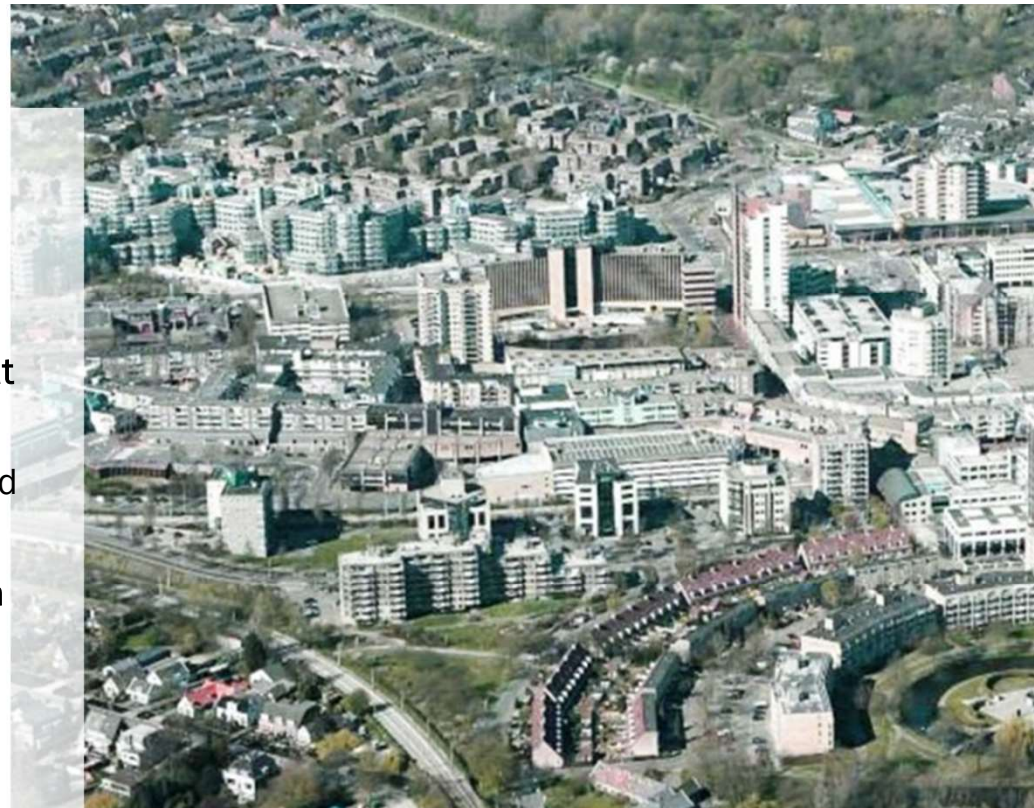
- › Data on social-emotional skills of youngsters may enhance the current policy model
- › Develop a hybrid policy model using theory and data and machine learning and 'traditional' statistics
- › Multidisciplinary team of data scientists, policy makers, legal advisors and domain experts

WVW	WVW	WVW	WVW	WVW
1. Kennisovername	2. Kennisovername	3. Kennisovername	4. Kennisovername	5. Kennisovername
6. Kennisovername	7. Kennisovername	8. Kennisovername	9. Kennisovername	10. Kennisovername
11. Kennisovername	12. Kennisovername	13. Kennisovername	14. Kennisovername	15. Kennisovername
16. Kennisovername	17. Kennisovername	18. Kennisovername	19. Kennisovername	20. Kennisovername
21. Kennisovername	22. Kennisovername	23. Kennisovername	24. Kennisovername	25. Kennisovername
26. Kennisovername	27. Kennisovername	28. Kennisovername	29. Kennisovername	30. Kennisovername
31. Kennisovername	32. Kennisovername	33. Kennisovername	34. Kennisovername	35. Kennisovername
36. Kennisovername	37. Kennisovername	38. Kennisovername	39. Kennisovername	40. Kennisovername
41. Kennisovername	42. Kennisovername	43. Kennisovername	44. Kennisovername	45. Kennisovername
46. Kennisovername	47. Kennisovername	48. Kennisovername	49. Kennisovername	50. Kennisovername
51. Kennisovername	52. Kennisovername	53. Kennisovername	54. Kennisovername	55. Kennisovername
56. Kennisovername	57. Kennisovername	58. Kennisovername	59. Kennisovername	60. Kennisovername
61. Kennisovername	62. Kennisovername	63. Kennisovername	64. Kennisovername	65. Kennisovername
66. Kennisovername	67. Kennisovername	68. Kennisovername	69. Kennisovername	70. Kennisovername
71. Kennisovername	72. Kennisovername	73. Kennisovername	74. Kennisovername	75. Kennisovername
76. Kennisovername	77. Kennisovername	78. Kennisovername	79. Kennisovername	80. Kennisovername
81. Kennisovername	82. Kennisovername	83. Kennisovername	84. Kennisovername	85. Kennisovername
86. Kennisovername	87. Kennisovername	88. Kennisovername	89. Kennisovername	90. Kennisovername
91. Kennisovername	92. Kennisovername	93. Kennisovername	94. Kennisovername	95. Kennisovername
96. Kennisovername	97. Kennisovername	98. Kennisovername	99. Kennisovername	100. Kennisovername



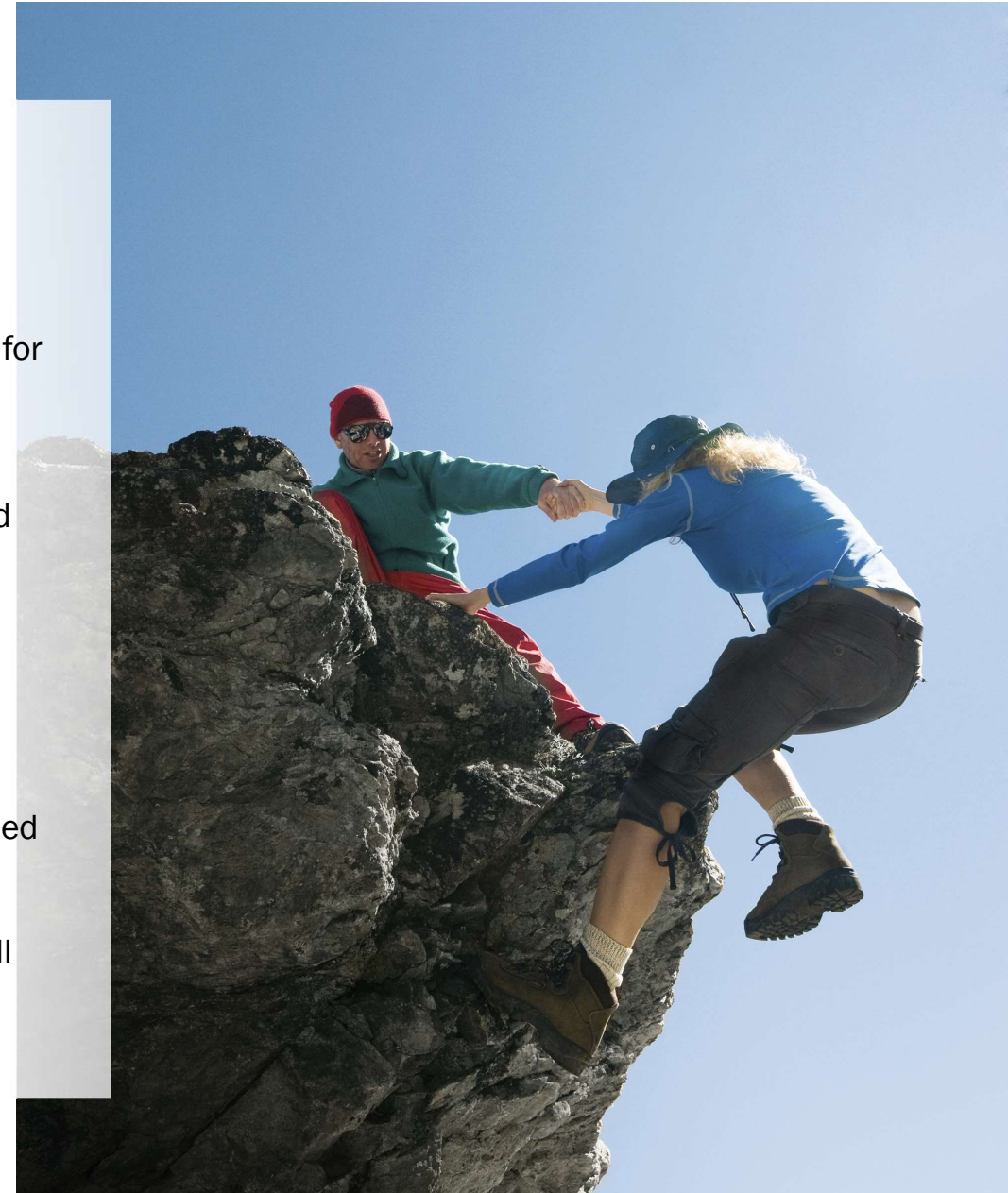
ENERGY TRANSITION MUNICIPALITY OF ZOETERMEER

- › Data experiment linking six **social challenges** to the **municipal heat transition** (natural gas free districts)
- › Analysis of new data sources to municipal information systems and analysis with **machine learning methods**
- › Multidisciplinary of the project is required to realize **collaboration between different departments** within the municipality: social domain, energy transition team and business intelligence / IT



› HOW TO SCALE?

- › Joint development of a **data processing agreement** is important for establishing trust
 - › ‘Top management support’ is essential
 - › Sometimes the GDPR remain unclear and may need to be refined
- › **Multidisciplinarity** is a requirement for useful outcomes
 - › Creating a joint understanding of the data and the model
- › Data is often created in a **different context** from where it is applied
 - › Data landscape often grew organically
 - › Interoperability; semantics and contextual understanding are all necessary for reuse of data
 - › Application of findings to new data may include biases



› IMPLICATIONS FOR POLICY

- › **Explainability of machine learning** is a challenge and may require new ways of openness
 - › Training the model is dependent on the data available
 - › A hybrid approach may include statistics testing the machine learning predictions
 - › In which way governments should become transparent: data, algorithm, process?
- › It is expected that the **policy cycle will accelerate**
 - › Different policy phases follow up more quickly and new links between phases emerge
 - › This also offers opportunities for rapid responses in case of undesired outcomes
 - › But new dependencies emerge: data is not always gathered and processed within the same organization



› **THANK YOU FOR
YOUR TIME**

DR. ANNE FLEUR VAN VEENSTRA
POLICY LAB LEAD

ANNEFLEUR.VANVEENSTRA@TNO.NL

**[HTTPS://WWW.TNO.NL/EN/FOCUS-AREAS/STRATEGIC-ANALYSIS-POLICY/EXPERTISE-
GROUPS/STRATEGY-POLICY/POLICY-LAB-DEVELOPING-DATA-DRIVEN-POLICIES/](https://www.tno.nl/en/focus-areas/strategic-analysis-policy/expertise-groups/strategy-policy/policy-lab-developing-data-driven-policies/)**

TNO innovation
for life